

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 8, 2005, 03:43:40 ; Search time 271.773 Seconds  
(without alignments)  
5783.472 Million cell updates/sec

Title: US-10-728-051-1  
Perfect score: 3562  
Sequence: 1 ataatacatatattcatc.....cgttgtgctgtttcttccc 2032

Scoring table: BLOSUM62

-Qs/cpn2 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Qs/cpn2 1/USPTO spool\_p/US10728051/runat\_07042005\_125044\_14041/app\_query.fasta\_1.4757  
-DB=A\_Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPFCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10728051 @CGN 1 1 672 @runat\_07042005\_125044\_14041 -NCPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq.i6Dec04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	92.3	626	2 AAY15244	Aay15244 Peanut al
2	3286	92.3	626	2 AAY25657	Aay25657 Peanut al
3	3286	92.3	626	6 ABUS2412	Abus2412 Peanut Ar
4	3286	92.3	626	7 ADC34906	Adc34906 Peanut al
5	3282	92.1	626	4 AAU04706	Aau04706 Anaphylac
6	3282	92.1	626	6 ABUS2570	Abus2570 Peanut Ar
7	3280	92.1	626	6 ABUS2568	Abus2568 Peanut Ar
8	3280	92.1	626	6 ABUS2571	Abus2571 Peanut Ar
9	3280	92.1	626	6 ABUS2574	Abus2574 Peanut Ar
10	3280	92.1	626	6 ABUS2573	Abus2573 Peanut Ar

11	3279	92.1	626	8 ADM12053	Adm12053 Arachis h
12	3278	92.0	626	6 ABUS2569	Abus2569 Peanut Ar
13	3278	92.0	626	6 ABUS2572	Abus2572 Peanut Ar
14	3274	91.9	626	2 AAW22150	Aaw22150 Peanut al
15	3269.5	91.8	625	7 ADG27464	Adg27464 Peanut Ar
16	3251	91.3	626	8 ADO38271	Ado38271 Peanut al
17	3052	85.7	634	3 AAB33599	Aab33599 Modified
18	3052	85.7	634	4 AAU04709	Aau04709 Modified
19	3052	85.7	634	4 AAU05034	Aau05034 Modified
20	3041	85.4	614	2 AAW22149	Aaw22149 Peanut al
21	3041	85.4	614	2 AAW62834	Aaw62834 Arachis h
22	3041	85.4	614	2 AAW52413	Aaw52413 Peanut Ar
23	3041	85.4	614	7 ADG27465	Adg27465 Peanut Ar
24	2130	59.8	415	2 AAY40913	Aay40913 Ara h 1 a
25	1281	36.0	605	2 AAW62838	Aaw62838 Glycine m
26	1281	36.0	605	7 ADG27564	Adg27564 Soybean B
27	1276	35.8	605	2 AAY40999	Aay40999 Soybean b
28	1190	33.4	559	5 ABG71270	Abg71270 Glycine m
29	1180.5	33.1	543	5 ABG71271	Abg71271 Glycine m
30	1165	32.7	417	5 ABG71269	Abg71269 Glycine m
31	1157	32.5	417	6 ABP97239	Abp97239 FLAG-tag
32	1157	32.5	425	6 ABP97238	Abp97238 Mature be
33	1156	32.5	439	6 ABP97235	Abp97235 Glycine m
34	1117.5	31.4	390	6 ABP97236	Abp97236 Glycine m
35	865.5	24.3	666	2 AAW62829	Aaw62829 Macadamia
36	863.5	24.2	625	2 AAW62830	Aaw62830 Macadamia
37	849.5	23.8	666	2 AAW62828	Aaw62828 Macadamia
38	811	22.8	566	2 AAR20181	Aar20181 Sequence
39	791	22.2	590	2 AAW62832	Aaw62832 Gossypium
40	764	21.4	371	2 AAY40914	Aay40914 Phaseolin
41	737.5	20.7	525	2 AAW62831	Aaw62831 Theobroma
42	654	18.4	366	6 ABUS2503	Abus2503 Soybean B
43	629.5	17.7	593	2 AAW62835	Aaw62835 Zea mays
44	626	17.6	573	7 ADL18493	Adl18493 Maize glo
45	614	17.2	637	2 AAW62837	Aaw62837 Hordeum v

ALIGNMENTS

RESULT 1  
AAY15244  
ID AAY15244 standard; protein; 626 AA.  
XX AC AAY15244;  
XX AC  
XX DT 17-OCT-2003 (revised)  
XX DT 09-NOV-1999 (first entry)  
XX DE Peanut allergen, Ara h 1, amino acid sequence.  
XX KW allergy; immune response; transgenic; allergen; epitope;  
XX KW immunoglobulin E; Ig E; binding site; peanut.  
XX OS Arachis hypogaea.  
XX PN WO9938978-A1.  
XX PD 05-AUG-1999.  
XX PF 29-JAN-1999; 99WO-US002031.  
XX PR 31-JAN-1998; 98US-0073283P.  
XX PR 13-FEB-1998; 98US-0074590P.  
XX PR 13-FEB-1998; 98US-0074624P.  
XX PR 17-FEB-1998; 98US-0074633P.  
XX PR 27-AUG-1998; 98US-00141220.  
XX PA (UYAR-) UNIV ARKANSAS.  
XX PA (UYNY ) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.  
XX PA (SOSI/) SOSIN H.  
XX PI Sosin H, Bannon GA, Burks AW, Sampson HA;  
XX

DR WPI; 1999-479189/40.  
 DR N-PSDB; AA206382.  
 XX Modified allergen with reduced IgE binding, useful for treating e.g.  
 PT allergies.  
 XX  
 PS Disclosure; Page 35-37; 46pp; English.  
 XX  
 CC This is the amino acid sequence of the Ara h 1 protein from Arachis  
 CC hypogaea. The Ara h 1 protein has 23 IgE (Immunoglobulin E) binding  
 CC epitopes, four of which are immunodominant (AAV15247, AAV15249, AAV15250  
 CC and AAV15263). By modifying the IgE binding sites the ability of the  
 CC allergen to provoke an immune response is downregulated. The epitopes of  
 CC the IgE binding sites can therefore be modified in genetically engineered  
 CC plants and animals to elicit less of an allergic response. (Updated on 17  
 CC -OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 626 AA;

Alignment Scores:  
 Pred. No.: 1,4e-304 Length: 626  
 Score: 3286.00 Matches: 626  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 92.25% Indels: 0  
 DB: 2 Gaps: 0

US-10-728-051-1 (1-2032) x AAV15244 (1-626)

QY 50 ATGAGAGGAGGTTCTCCACATGATGCTGTGCTAGGATCCTGTCTGCTCAGTT 109  
 DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20  
 QY 110 TCTGAAGCGATGCAAGTCATCACCTACCAAGAGAAACAGAGAACCCCTGCCCGAC 169  
 DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40  
 QY 170 AGGTGCTCCAGATTGTCAACAGGAACCGGATGACTTGAAGCAAAAGSCATGCGAGTCT 229  
 DB 41 ArgCysLeuGlnSerCysGlnGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60  
 QY 230 CGCTGCACCAAGCTCGAGTATGATCCTGTTGTGTCTATGATCCTCGAGGACACACTGGC 289  
 DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80  
 QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGACACAGTGGCGGCCCAACCCGAGACTAC 349  
 DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100  
 QY 350 GATGATGACCGCGCTCAACCCGAGAGAGGAGGCGCGATGGGACCACTGGACCG 409  
 DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTyrProAlaGlyPro 120  
 QY 410 AGGGAGCTGAAGAGAGAAAGACTGGAGACAACCAAGAGAGATTGGAGGCGACCAAGT 469  
 DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140  
 QY 470 CATCAGCAGCCAGGAAAATAAGGCCGAGGAAGAGAGAGAAACAAGAGTGGGGAACA 529  
 DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnIleArgTyr 160  
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGAAACAACCCCTTCTACTTCCCGTCAAGG 589  
 DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180  
 QY 590 CGGTTTAGCCCGCTACGGGACCAACACCGTAGGATCCGGTCTCGAGGTTTGCAC 649  
 DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgGileArgValLeuGlnArgPheAsp 200  
 QY 650 CAAAGTCAAGGCAAGTTTCAGATCTCCAGAAATCACCGTATTGTGTGAGATCCAGGCCAAA 709  
 DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220

QY 710 CTTAACACTCTGTGTCTTCCCAAGCAGCTGATGCTGATAAATCATCTTGTATTCACAGAA 769  
 DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240  
 QY 770 GGGCAGCCACCGTACCGTAGCAATAATGGCAATAACAGAAAGAGCTTTAACTTTCACGAG 829  
 DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260  
 QY 830 GGCCATGCACCTCAGAAATCCCATCCGTTTCTATCTTCTACATCTTGAACCGCATCACAAC 889  
 DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280  
 QY 890 CAGAACTCAGAGTAGCTAAATCTTCATGCCGCTTAACACACCCGCGCAGCTTTCAGAT 949  
 DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300  
 QY 950 TTCTTCCCGGAGGAGCGGAGACCAATCATCTTCTGAGGCGCTTTCAGAGGATACG 1009  
 DB 301 PhePheProAlaSerSerArgAspGlnSerTyrLeuGlnGlyPheSerArgAsnThr 320  
 QY 1010 TTGGAGCGCGCTTCAATGCGGAATTCATCAGATACCGAGGCTGCTGTAGAACAGAAAT 1069  
 DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340  
 QY 1070 GCAGAGGTGAGCAAGAGAGAGGAGGCGAGGATGAGTACTCGAGTAGTGAGAAC 1129  
 DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360  
 QY 1130 AATGAGAGCTGATAGTCAAAAGTGTCAAGAGGAGCGCTTGAAGACTTACTTAAGCACCT 1189  
 DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380  
 QY 1190 AAATCCGTCTCAAGAAAGGCTCCGAAAGAGGAGGAGATATCAACCAACCAATCAACTTG 1249  
 DB 381 LysSerValSerLysLysGlySerGluGluGluGluGlyAspIleThrAsnProIleAsnLeu 400  
 QY 1250 AGAGAGGCGAGCGCGATCTTTTAAACAACTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309  
 DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420  
 QY 1310 AAGAGAACCCCGCTCAGAGCTGACATGATGCTCAGCTGTGTAGAGATCAAGAA 1369  
 DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluLysGlu 440  
 QY 1370 GGAGCTTTGATGCTCCCACTTCAACTCAAGGCGCATGTTATCGTCTGCTCAACAAA 1429  
 DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460  
 QY 1430 GGAACTGGAACCTTGAACCTGCTGCTGTAAAGAAAGAGCAACACAGAGGCGGCGCG 1489  
 DB 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 480  
 QY 1490 GAAGAAGAGGAGGAGCGAAGAGAGAGAGGAGGAGTAACAGAGAGTGGTGGTAGGTAC 1549  
 DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500  
 QY 1550 ACAGCGAGTTTGAAGAAAGGCGATGTTTCTATGCTCCAGCAGAGCTCATCCAGTAGCCATC 1609  
 DB 501 ThrAlaArgLeuLysGluLysValPheIleMetProAlaAlaHisProValAlaIle 520  
 QY 1610 AACGTTCTCCGAACTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1669  
 DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540  
 QY 1670 ATCTTCTTGGAGGTGATAAGGACAAATGTGTATAGACACAGATAGAGAAGCAAGCGAAGAT 1729  
 DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560  
 QY 1730 TTAGCATTCCTTGGGTCCGGTGAAACAAGTTGAGAGCTCATCAAAAACAGAAAGAAATCT 1789  
 DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580  
 QY 1790 CACTTTGTGAGTGTGCTCGTCCCTCAATCTCAATCTCAATCTCGTCTCGTCTCGAGAAAGAG 1849



Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340  
QY 1070 GCAGGAGGTGACCAAGAGGAGGCGCAGAGCGATGAGTACTCGAGTAGTCAGAAC 1129  
Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgGlySerThrArgSerSerGluAsn 360  
QY 1130 AATGAAGGAGTGTAGTCAAAAGTGTCACAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCT 1189  
Db 361 AsnGluGlyValIleValIleValSerLysGluHisValGluGluLeuThrLysHisAla 380  
QY 1190 AAATCCGTCTCAAAAGAGGCTCCGAAGAGGAGGAGATATCACCAACCAATCAACTTG 1249  
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400  
QY 1250 AGNAGGCGAGCCCATCTTCTTAACACTTGGGAAGTTATTCAGGTGAAGCCAGAC 1309  
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnGlyLysLeuPheGluValLysProAsp 420  
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGGACATGCTCACCTGTGTAGAGATCAAGAA 1369  
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440  
QY 1370 GAGCTTTGATGTCCTCCACTCAACTCAAGGCCATGGTTATCTGCTGTCACAA 1429  
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460  
QY 1430 GGAAGTGGAACTTGAACCTGCGCTGTAAAGAAAGAGCAACACAGAGGCGCGG 1489  
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480  
QY 1490 GAAGAAGAGGAGGACGAAGACGAAGAGGAGGAGGAGTAAACAGAGGTCGTAGGTAC 1549  
Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500  
QY 1550 ACAGCGAGGTGAAGAGCGGATGTCTCATCATGCCAGCAGCTATCCAGTAGCCATC 1609  
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520  
QY 1610 AACGTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACCTGAAACACACAGA 1669  
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540  
QY 1670 ATCTTCTTCAGGTGATAGGACAATGTGATAGACAGATAGACGAAGCGAAGGAT 1729  
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560  
QY 1730 TTAGCATTCCTGGTCCGGTGACAGTTGAGAGCTCATCAAAACACAGAGGAATCT 1789  
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580  
QY 1790 CACTTTGTGAGTGTCTCCTCAATCTCAATCTCAATCTCCGTCGTCTCCTGAGAAAGAG 1849  
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600  
QY 1850 TCTCCTGAGAAAGAGATCAAGAGGAGGAAACCAAGAGGAGGAGGTCCTCCTTTCA 1909  
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620  
QY 1910 ATTTTGAAGCTTTTAC 1927  
Db 621 IleLeuLysAlaPheAsn 626

## RESULT 3

ABU52412  
ID ABU52412 standard; protein; 626 AA.  
XX  
AC ABU52412;  
XX  
DT 10-MAR-2003 (first entry)  
XX  
DE Peanut Ara h1 protein from cDNA clone p41b.  
XX  
KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IGE binding site;

anaphylactic food allergen; anti-allergenic; vaccine; wound healing.  
XX  
OS Arachis hypogaea.  
XX  
PN WO200274250-A2.  
XX  
PD 26-SEP-2002.  
XX  
PF 18-MAR-2002; 2002WO-US009108.  
XX  
PR 16-MAR-2001; 2001US-0276822P.  
XX  
PR 18-MAR-2002; 2002US-00276822.  
XX  
PA (PANA-) PANACEA PHARM.  
XX  
PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;  
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;  
PI Rabin PA, Shin DS, Stanley JS;  
XX  
DR WPI; 2003-018765/01.  
XX  
DR N-PSDB; ABX70603.  
XX  
PT New modified anaphylactic food allergen, useful for preventing or  
PT treating allergic reactions associated with e.g. anaphylactic allergens.  
XX  
PS Claim 27; Fig 13; 300pp; English.  
XX  
CC The invention relates to a modified anaphylactic food allergen has an  
CC amino acid sequence that is substantially identical to that of natural  
CC anaphylactic food allergen, except for a cysteine residue that has been  
CC modified so that it cannot participate in the disulphide bond. The  
CC modification may also comprise mutation of the IGE binding sites to  
CC reduce allergenicity. Also included are: (1) a method of making a  
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding  
CC or for causing a site specific mutation in the modified anaphylactic food  
CC allergen; (3) a transgenic plant or animal expressing the modified  
CC anaphylactic food allergen; (4) a method of treating an individual by  
CC reducing the clinical response to a natural anaphylactic food allergen;  
CC and an isolated fragment of peanut allergen Ara h 1. The modified  
CC anaphylactic food allergen is useful for preventing or treating allergic  
CC reactions associated with any natural allergen such as food, insect,  
CC rubber or preferably anaphylactic allergens. It is also useful for  
CC treating wounds in mammals such as bovine, canine, feline, caprine,  
CC ovine, porcine, murine or equine species. The present sequence is a  
CC peanut allergen (e.g. Ara h1, h2 or h3)  
XX  
SQ Sequence 626 AA;

## Alignment Scores:

Pred. No.: 1.4e-304 Length: 626  
Score: 3286.00 Matches: 626  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.25% Indels: 0  
Gaps: 0

US-10-728-051-1 (1-2032) x ABU52412 (1-626)

QY 50 ATGAGAGGAGGGTTCCTCCACTGATGCTGTCTAGGATCCTGTCTCGCTTCAGTT 109  
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20  
QY 110 TTGCAACCGATGCCAGTCATCACCTTACCAAGAGAAACAGAGAACCCCTGCCGCCAG 169  
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40  
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAAACCGGATGACTTCAAGCAAGGATCGAGTCT 229  
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60  
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTATGATCTCTCGAGGACACTGGC 289  
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80



Qy	290	ACCAACCAACGTTTCCCTCTCAGGGAGCGGACACGTGGCCGCCAACCCGGAGACTTAC	349
Db	81	ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
Qy	350	GATCATGACCGCTCAACCCGGAAGAGAGGAGCGCCGATGGGACCAGCTGGACCG	409
Db	101	AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro	120
Qy	410	AGGAGCGCTGAAGAGAGAAGACTGGAGACAACCAAGAGAAGATTGGAGCGCACCAAGT	469
Db	121	ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer	140
Qy	470	CATCAGCAGCCACGGAAATAAGGCCCGAAGAAAGAGAGAGAAACAAGAGTGGGGAAACA	529
Db	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTyrGlyThr	160
Qy	530	CCAGGTAGCCATGATGAGGAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGG	589
Db	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
Qy	590	CGGTTTAGCACCCCTACTCGGAAACCAAAACGGTAGGATCCGGGTCTCGACAGGTTTGAC	649
Db	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
Qy	650	CAAAAGTCAAGCGAGTTTCAGAACTCCAGAAATCACCGTATGTGTCAGATCGAGGCCAAA	709
Db	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
Qy	710	CCTTAACACTCTTGTTCTTCCCAACAGCCTGATCGCTGATACATCTCTTGTTATCCACAA	769
Db	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
Qy	770	GGGCAAGCCACCGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG	829
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
Qy	830	GGCCATGCACCTCAGAAATCCCATCCCGTTTCAATTTCTACATCTTTGAACCCGCATGACAAC	889
Db	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
Qy	890	CAGAACCTCAGAGTAGCTAAATCTCCATGCCCGCTTAACACACCCGCGCAGTTTGAGAT	949
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
Qy	950	TTCTTCCGGCGGACGACGCGAGACCAATCATCTTACTTGCAGGSGCTTCAGCAGGAATACG	1009
Db	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
Qy	1010	TTGAGGCGCGCTTCAATGCGGAATTCAATGAGATACGGAGGTGCTGTTTAGAGAGAAT	1069
Db	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	340
Qy	1070	GCAGAGGTGAGCAAGAGAGAGAGAGGCGAGGCGATGGAGTACTCGGAGTAGTGAGAAC	1129
Db	341	AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrPheSerSerGluAsn	360
Qy	1130	AATCAAGGAGGTAGTCAAGTCTCAAGAGAGACAGCTTGGAAGACTTACTTAAGCAGCT	1189
Db	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
Qy	1190	AAATCCGTCTCAAAAGAAAGGCTCCGAAAGAGAGGAGATATCAACCAACCCCAATCAACTTG	1249
Db	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
Qy	1250	AGAGAAGCGGACCCGATCTTTTCAACAACCTTGGGAAGTTATTGTAGGTGAAGCCAGAC	1309
Db	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
Qy	1310	AAGAAGRACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTCTGTAGAGATCAAGAA	1369
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440

QY	1370	GGAGCTTTGATGCTCCACACTTCAACTAAAGGCCATATGGTTATTCGTCGTCGCAACAA	1429
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAACCTCGAACCCTTGAACTTCGTGGCTGTGAAGAAAGAGCAACAAACAGAGGGGACGGCGG	1489
Db	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgLysArg	480
QY	1490	GAAGAAGAGGAGGACGAAGACGAAGAAGAGAGGAGGGAAGTAACAGAGAGGTGCGTAGGTAC	1549
Db	481	GluGluGluGluAspGluAspGluGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACACGAGGTTGAAGGAAGCGGATGTTTCATCATGTCAGAGGCTCATCCAGTAGCCATC	1609
Db	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisProValAlaIle	520
QY	1610	AACGCTTCCTCCGAACCTCCATCTGCTGGCTTCGGTATCAACGCTGAAAAACAACACACAGA	1669
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCTCTGCAGGTGATAAGGACAATGTGTATAGACCAGATAGAGAAAGCAAGCGAAGGAT	1729
Db	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTCCTCGGTCGGGTGAACAAGTTGAGAGAGCTCATCAAAAAACCAAGAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACTTTCGTAGTGCTCTCTCTCAATCTCAATCTCCGTCGCTCTCTGAGAAAGAG	1849
Db	581	HisPheValSerAlaArgProGlnSerGlnSerProSerSerProGluLysGlu	600
QY	1850	TCTCCTCAGAAAAAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGTCCTCACTCTCTTCA	1909
Db	601	SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGCTTTTAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
RESULT 4			
ADC34906			
ID	ADC34906 standard; protein; 626 AA.		
XX	ADC34906;		
AC	ADC34906;		
DT	18-DEC-2003 (first entry)		
XX	Peanut allergen Ara h 1.		
DE	allergen; antigen; hyporesponsive; desensitisation; immunomodulator;		
KW	gene therapy; peanut.		
XX	Arachis hypogaea.		
OS	WO2003047618-A2.		
FN	12-JUN-2003.		
XX	05-DEC-2002; 2002WO-GB005548.		
PF	05-DEC-2001; 2001US-0338385P.		
XX	(CIRC-) CIRCASSIA LTD.		
PA	Larche M, Ledger PW;		
XX	WPI; 2003-523267/49.		
DR	Desensitizing an individual to a selected polypeptide antigen comprises		
PT	administering a composition containing polypeptide antigens in an amount		
PT	that generates a state of hyporesponsiveness to the antigen to allow		
PT	desensitization.		

Desensitizing an individual to a selected polypeptide antigen comprises administering a composition containing polypeptide antigens in an amount that generates a state of hyporesponsiveness to the antigen to allow desensitization.

XX PS Disclosure; Page 40; 57pp; English.

XX CC The invention relates to a novel method for desensitising an individual  
 CC to a selected polypeptide antigen. The method comprises administering a  
 CC composition that contains polypeptide antigens in an amount that  
 CC generates in the individual a state of hyporesponsiveness to the antigen  
 CC to allow desensitisation to one or more polypeptide antigens. The method  
 CC of the invention has immunomodulator activity, and may have a use in gene  
 CC therapy. The composition and method are useful in manufacturing a  
 CC medicament for desensitising an individual to a selected polypeptide  
 CC antigen or for generating in the individual a state of hyporesponsiveness  
 CC to the antigen to allow desensitisation to one or more polypeptide  
 CC antigens. The present sequence is used in the exemplification of the  
 CC invention.

XX CC Sequence 626 AA;

XX SQ

Alignment Scores:

Pred. No.:	1.4e-304	Length:	626
Score:	3286.00	Matches:	626
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	92.25%	Indels:	0
DB:	7	Gaps:	0

US-10-728-051-1 (1-2032) x ADC34906 (1-626)

QY 50 ATGAGAGGGAGGTTTCTCCACGTGATGCTGTTGCTAGGAGATCCTTGCTCGGTTTCAGTT 109

DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyLeuValLeuAlaSerVal 20

QY 110 TCTGCAAGCATGCCAAGTTCATCACCCTTACAGAGAAACACAGAACCCCTCGGCCAG 169

DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

QY 170 AGTGCCCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGGAGTCT 229

DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60

QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTGTCTATGATCCTCGAGCACACACTGGC 289

DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGCGGACACAGTGGCCGCCCAACCCGGAGACTAC 349

DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100

QY 350 GATGATGACCCCGTCAACCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 409

DB 101 AspAspAspArgArgGlnProArgGlnGluGlyArgGlyArgGlyArgGlyArgGlyPro 120

QY 410 AGGGAGCGTGAAG 469

DB 121 ArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGlu 140

QY 470 CATCAGCAGCCAGCAATGAAGCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529

DB 141 HisGlnGlnProArgLysLeuArgProGluGlyArgGlyArgGlyArgGlyArgGlyThr 160

QY 530 CCAGGTACCCCGTCAAG 589

DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180

QY 590 CGGTTTACCCCGTCAAG 649

DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgLeuValLeuGlnArgPheAsp 200

QY 650 CAACGGTCAAGCAGTTTCAGAAATCTCCAGAAATCAACCGTATTTGTCAGATCGAGCCCAA 709

DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgGlyLeuGlnGlnGlnGlnGln 720

QY 710 CCTACACACTCTGTTCTTCCCAAGCAGCGCTGATGATGATGATGATGATGATGATGATGAT 769

DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240

QY 770 GGGCAAGCCACCGTACCGTACCAATGGCAATAGCAATAACAGAAAGAGCTTTAATCTTACGAG 829

DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260

QY 830 GGCCATGCATCTCAGAAATCCCATCCGTTTCAATTTCTTACATCTTTGAACCCCGCATGACAAC 889

DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280

QY 890 CAGAACCTCAGACTAGCTAAATCTCCATGCGCTTACACACACCGCCGCTTGTAGGAT 949

DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300

QY 950 TTCTTCCCGCGAGCAGCGCAGAGCAATCATCTTCTTACCTTTCAGGCTTCAGCAGGAATACG 1009

DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320

QY 1010 TTGGAGCGCGCTTCAATGCGGAATTCATGAGATACGGAGGGTGTCTGTAGAGAGNAT 1069

DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340

QY 1070 GCAGGAGCTGACGAG 1129

DB 341 AlaGlyGlyGluGlnGlnGluArgGlyGlnArgArgIlePheSerThrArgSerSerGluAsn 360

QY 1130 AATGAAGGAGTATAGTCAAAAGTGTCAAAAGGAGCAGCTTGAAGAACTTCTTACGACGCT 1189

DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380

QY 1190 AATCGCTCTCAAGAAAGGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1249

DB 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProLleAsnLeu 400

QY 1250 ACAGAGAGCGAGCCCGATCTTCTTAACTTGGAAAGTATTATTTAGGTGAAGCCAGAC 1309

DB 401 ArgGluGlyGluProAspLeuSerAsnAsnAsnPheGlyLysLeuPheGluValLysProAsp 420

QY 1310 AAGAAGAACCCCGCTTCCAGACTTCCAGACTGATGATGCTCCTGCTGAGAGTCAAGAA 1369

DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 440

QY 1370 GGAGCTTTCATCTCCACCTTCACTCAAGGCGATGTTTCTGCTGCTGCTCAACAAA 1429

DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460

QY 1430 GGAACCTGGAAACCTTGAACCTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGAGCGCGG 1489

DB 461 GlyThrGlyAsnLeuGluLeuValAlaValA-gLysGluGlnGlnArgGlyArgArg 480

QY 1490 GAAGAAG 1549

DB 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500

QY 1550 ACAGCCAGGTGTGAAGAGAGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609

DB 501 ThrAlaArgLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520

QY 1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAG 1669

DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540

QY 1670 ATCTTCTTCAGGTGATAGGACATGTGATAGACAGATAGAGAGAGAGAGAGAGAGAGAGAT 1729

DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560

QY 1730 TTAGCATTTCCCTGGGTGGGTGAACCAAGTTTGAGAACTCATCAAAACCAAGAGGAATCT 1789

DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580

QY 1790 CACTTGTGAGTCTGCTCTCAATCTCAATCTCAATCTCCGTCGCTCTCTCAGAAAGAG 1849

Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600  
QY 1850 TCTCTCTGAGAAAGAGATCAAGAGGAGGAAACCAAGAGGAGGATCCACTCTTTCA 1909  
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyGlyGlyProLeuLeuSer 620  
QY 1910 ATTTTGAAGGCTTTTAAAC 1927  
Db 621 IleLeuLysAlaPheAsn 626

RESULT 5  
AAU04706  
ID AAU04706 standard; protein; 626 AA.  
XX AAU04706;  
AC AAU04706;  
XX 23-OCT-2001 (first entry)  
DE Anaphylactic antigen Ara h 1.  
XX Ara h 1; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;  
KW allergy; mast cell; basophil; mouse.  
XX Mus sp.

XX Key Location/Qualifiers  
PH Misc-difference 285 /note= "Encoded by gta"  
FT FT  
XX WO200140264-A2.  
XX 07-JUN-2001.  
XX 06-DEC-2000; 2000MO-US033124.  
XX 06-DEC-1999; 99US-00455294.  
XX 23-JUN-2000; 2000US-0213765P.  
XX 27-SEP-2000; 2000US-0235797P.  
XX (PANA-) PANACEA PHARM LLC.  
PA (UYAR-) UNIV ARKANSAS.  
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.  
XX Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;  
XX WPI; 2001-381378/40.  
XX N-PSDB; AAS08537.  
XX Antigenic fragments useful for reducing anaphylactic risk and reducing  
PT the severity and/or number of allergic symptoms in individuals sensitive  
PT to antigens, have reduced ability to bind Immunoglobulin E.  
XX Claim 7; Fig 9; 100pp; English.

XX The sequence represents the amino acid sequence of anaphylactic antigen  
CC Ara h 1. Ara h 1 is an anaphylactic antigen (A), which was used to design  
CC antigenic peptides having a reduced ability to bind immunoglobulin E  
CC (Ige) as compared with the intact (A), or having a sequence substantially  
CC identical to a portion of sequence of an antigen that includes at least  
CC one Ige binding site, where at least one Ige binding site of the peptide  
CC is altered. The antigenic peptides are used in a composition which is  
CC useful for reducing risk or severity of allergic reaction to an antigen.  
CC This is done by identifying an individual at risk of allergic reaction to  
CC an antigen by identifying prior display of allergic symptoms when exposed  
CC to the antigen, or a familial relationship with an individual who  
CC previously displayed allergic symptoms when exposed to the antigen.  
CC Following this an antigen-specific Ige present on one or more mast cells  
CC or basophils in the individual's serum is identified. The individual is  
CC then contacted with a peptide corresponding to a portion of the antigen,  
CC which is selected, formulated, and delivered so that binding of the  
CC peptide to antigen-specific Ige is reduced as compared with Ige binding  
CC of intact antigen. The composition is also useful for treating and  
CC preventing allergic reactions

XX SQ Sequence 626 AA;  
Alignment Scores:  
Pred. No.: 3,38e-304 Length: 626  
Score: 3282.00 Matches: 625  
Percent Similarity: 99.84% Conservatives: 0  
Best Local Similarity: 99.84% Mismatches: 1  
Query Match: 92.14% Indels: 0  
DB: 4 Gaps: 0

US-10-728-051-1 (1-2032) x AAU04706 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCAGTGTCTGCTAGGATCCTTGTCTGGTTCAGTT 109  
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20  
QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAAAGAAACAGAACCCCTCGGCCAG 169  
Db 21 SerAlaThrHisAlaLysSerProTyrGlnLysThrGluAsnProCysAlaGln 40  
QY 170 AGGTGCTCCAGAGTTGTCAACAGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229  
Db 41 ArgCysLeuGlnSerCysGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60  
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACATGGC 289  
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80  
QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGGACACGTCGGCCGCCAACCCGGAGACTAC 349  
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100  
QY 350 GATGATGACCGCGCTCAACCCCAAGAGAGGAGGCGGATGCGGACCGAGCTGGACCG 409  
Db 101 AspAspAspArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120  
QY 410 AGGAGCGTGAAAGAGAGAAAGACTGGAGACCAACCAAGAGAAAGATTGGAGCGCACCAAGT 469  
Db 121 ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgGpCoser 140  
QY 470 CATCAGCAGCCCGGAAATTAAGCCCGAGGAGAGAGGAGAAAGAGCAAGAGTCGGGACACA 529  
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluLutPFGlyThr 160  
QY 530 CCAGGTAGCCATCTGAGGGAAGAAAACATCTCGGAACAACCTTTCTACTTCCCGTCAAGG 589  
Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180  
QY 590 CGGTTTAGCACCGCTTACCGGAAACCAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGAC 649  
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200  
QY 650 CAAAGGTCAAGCGAGTTTCAGATCTCCAGATCCAGATCCAGTATTTGTCAGATCCAGGCCAAA 709  
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220  
QY 710 CTTAAACACTCTTGTCTTCTCCCAAGCACGCTGATGCTGATAAATCTTGTATTCCAGCAA 769  
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240  
QY 770 GGGCAAGCCACCGTGACCGTAGCAATGGCAATTAACAGAAAGAGCTTTAATCTTGACGAG 829  
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260  
QY 830 GGCCATGCACTCAGATCCATCCCGGTTTCATTTCTCTACATCTTGAACCGCCATGACAC 889  
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280  
QY 890 CAGAACTCAGAGTAGTAAATCTCCATGCCCGTTAAACACACCCCGGAGTTTGAGGAT 949  
Db 281 GlnAsnLeuArgAlaAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300

QY 950 TTCTCCCGCGGAGGAGCGGAGACCAATCATCTTCTGAGGGCTTCAGCAGGAATACG 1009  
 Db |||||  
 QY 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320  
 Db |||||  
 QY 1010 TTGGAGCGCGCTTCAATCGGAATTCATAGATACGAGGGTCTGTTAGAGAGAAAT 1069  
 Db |||||  
 QY 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340  
 QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGAGGCGAGCGGAGTGCAGTACTCGGAGTACTGAGAAC 1129  
 Db |||||  
 QY 341 AlaGlyGlyGlnGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360  
 QY 1130 AATGAAGGAGTGTAGTCAAGTGTCAAGGAGCACGTTGAAGAACTTACTTAAGCACGCT 1189  
 Db |||||  
 QY 361 AsnGluGlyValIleValIleValIleValIleValIleValIleValIleValIleValIle 380  
 QY 1190 AATCCGTCTCAAGAAAGAGCTCCGAGAGAGGAGATATACCAACCAATCAATCACTTG 1249  
 Db |||||  
 QY 381 LysSerValSerLysLysGlySerGluGluGluGluGluGluGluGluGluGluGluGlu 400  
 QY 1250 AGAAGGCGGAGCGGATCTTCTCAACACTTTGGGAGGTATTTGAGGTGAAGCCAGAC 1309  
 Db |||||  
 QY 401 ArgGluGlyGluProAspLysSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420  
 QY 1310 AAGAAGAACCCAGCTTTCAGGACCTGGACATGCTCACCTGTGTAGAGATCAAAAGAA 1369  
 Db |||||  
 QY 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440  
 QY 1370 GGAGCTTTGATGCTCCACACTCAAGGCGATGTTATCTCTGCTCAACAA 1429  
 Db |||||  
 QY 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValValValVal 460  
 QY 1430 GGAACTCGAAACCTTGAACTCGTGGCTGTAAAGAAAGAGCAACAGAGGGGAGCGCGG 1489  
 Db |||||  
 QY 461 GlyThrGlyAsnLeuGluValAlaValAlaValArgLysGluGlnGlnArgGlyArgArg 480  
 QY 1490 GAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1549  
 Db |||||  
 QY 481 Glu 500  
 QY 1550 ACAGCGAGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1609  
 Db |||||  
 QY 501 ThrAlaArgLeuLysGluGlyAspValPheLysMetProAlaAlaHisProValAlaIle 520  
 QY 1610 AACGCTTCTCCGAACTCCATCTCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCA 1669  
 Db |||||  
 QY 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540  
 QY 1670 ATCTTCTGAGGTGATAGGACAATGTGTAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1729  
 Db |||||  
 QY 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560  
 QY 1730 TTAGCATTCCTGCTGGTGGGTGAACAACTTCAGAAAGCTCATCAAAACCAAGAGGAATCT 1789  
 Db |||||  
 QY 561 LeuAlaPheProGlySerGlyGlnValGluLysLeuIleLysAsnGlnLysGluSer 580  
 QY 1790 CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTCAAGAGAG 1849  
 Db |||||  
 QY 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600  
 QY 1850 TCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909  
 Db |||||  
 QY 601 SerProGluLysGluAspGlnGluGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 620  
 QY 1910 ATTTTGAAGGCTTTTAAAC 1927  
 Db |||||  
 QY 621 IleLeuLysAlaPheAsn 626  
 RESULT 6  
 ID ABUS2570  
 AC ABUS2570 standard; protein; 626 AA.

XX 10-MAR-2003 (first entry)  
 XX Peanut Ara h1 mutant V72A.  
 DE Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;  
 XX mutein; anaphylactic food allergen; anti-allergenic; vaccine;  
 KW wound healing.  
 XX Homo sapiens.  
 OS WO200274250-A2.  
 XX 26-SEP-2002.  
 XX 18-MAR-2002; 2002WO-US009108.  
 XX 16-MAR-2001; 2001US-0276822P.  
 XX 18-MAR-2002; 2002US-00276822.  
 XX (PANA-) PANACEA PHARM.  
 PA Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;  
 PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;  
 PI Rabin JM, Shin DS, Stanley JS;  
 XX WPI; 2003-018765/01.  
 XX New modified anaphylactic food allergen, useful for preventing or  
 PT treating allergic reactions associated with e.g. anaphylactic allergens.  
 XX Example 5; Page; 300pp; English.  
 CC The invention relates to a modified anaphylactic food allergen has an  
 CC amino acid sequence that is substantially identical to that of natural  
 CC anaphylactic food allergen, except for a cysteine residue that has been  
 CC modified so that it cannot participate in the disulphide bond. The  
 CC modification may also comprise mutation of the IgE binding sites to  
 CC reduce allergenicity. Also included are: (1) a method of making a  
 CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding  
 CC or for causing a site specific mutation in the modified anaphylactic food  
 CC allergen; (3) a transgenic plant or animal expressing the modified  
 CC anaphylactic food allergen; (4) a method of treating an individual by  
 CC reducing the clinical response to a natural anaphylactic food allergen;  
 CC and an isolated fragment of peanut allergen Ara h 1. The modified  
 CC anaphylactic food allergen is useful for preventing or treating allergic  
 CC reactions associated with any natural allergen such as food, insect,  
 CC rubber or preferably anaphylactic allergens. It is also useful for  
 CC treating wounds in mammals such as bovine, canine, feline, caprine,  
 CC ovine, porcine, murine or equine species. The present sequence is a food  
 CC allergen, mutated to alter its IgE binding characteristics. Note: The  
 CC present sequence is not shown in the specification but was created by the  
 CC indexer using information provided in the specification  
 XX SQ Sequence 626 AA;  
 Alignment Scores:  
 Pred. No.: 3,38e-304 Length: 626  
 Score: 3282.00 Matches: 625  
 Percent Similarity: 99.84% Conservative: 0  
 Best Local Similarity: 99.84% Mismatches: 1  
 Query Match: 92.14% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-728-051-1 (1-2032) x ABUS2570 (1-626)  
 QY 50 ATCAGAGGAGGGTTCCTCCACTGATGCTGCTAGGATCCTGCTGCTCAGTT 109  
 Db |||||  
 QY 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuLeuValLeuAlaSerVal 20  
 QY 110 TCTGCAACGATCCAGTTCATCCTTACCAAGAAAAACAGAAACCCCTGCGCCAG 169  
 Db |||||  
 QY 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysThrGluAsnProCysAlaGln 40

QY	170	AGGTGCTCCAGATTGTCACACAGGAAACCGGATGACTTGAGCAAAAGGACATCGAGTCT	1250	AGAGAGGCGAGCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGAC	1309
DB	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	230	CGCTGCACCAAGCTCCAGATGATGATCTCGTTGTCTATGATCCTCGAGGACACACTGGC	1310	AAGAGAAACCCAGCTTCAGGACCTGACATGATCTCACCTGTGTAGAGATCAAGAAG	1369
DB	61	ArgCysThrLysLeuGluLysAspProArgCysAlaLysAspProArgGlyHisThrGly	421	LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluLysGlu	440
QY	290	ACCACCAACCAAGCTTCCCTCCAGGGAGCGACACGTCGGCCGCAACCCGAGAGACTAC	1370	GGAGCTTTGATGCTCCACACTTCACTCAAAAGGCCATGGTTATCTCGTCGTCAACAA	1429
DB	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	350	GATGATGACCGCGCTCAACCCGAAAGAGAGAGAGCGCGATGGGACAGCTGACCG	1430	GGAACTGGAAACCTTGAACCTCGTGGCTGTAAAGAAAGACAAACAAGAGGGGACGGCG	1489
DB	101	AspAspArgArgGlnProArgGluGluGlyGlyArgTrpGlyProAlaGlyPro	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg	480
QY	410	AGGAGCGTGAAGACAGAAGACTCGAGACACCAAGAGAGATGGAGGCGACCACT	1490	GAAGAGAGGAGGACGACGACCAAGAGAGGAGGAGTAACAGAGAGGTGGCTAGGTAC	1549
DB	121	ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer	481	GluGluGluGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	470	CATCAGCAGCACGGGAAATAAGCCCGAGAGAGAGAGAAACAAGAGTGGGAAACA	1550	ACAGCGAGGTTGAAGGAGCGGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
DB	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnTrpGlyThr	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	530	CMAGTAGCCATGTGAGGAGAAACAATCTCGGAAACAACCTTTCTATTCCGCTCAAG	1610	AACGCTTCTCCGAACCTCCATCTCGTTCGCTTCGCTATCAACGCTGAAACCAACAG	1669
DB	161	ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg	540
QY	590	CGGTTTAGCACCGCTACCGGACCAACAAACGGTAGGATCGGCTCTCGAGGTTGAC	1670	ATCTTCTTGCAGGTGATAGGACCAATGTAGTACAGCAGATAGAGAAAGCAAGCAAGAT	1729
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	650	CAAAGCTCAAGGAGGTTTCAAGATCTCCAGAATCACCGTATTGTGAGATCGAGGCCAA	1730	TTAGCATTCCTCGGCTCGGCTGAAAGTTGAGAACTCATCAAAAACCCAGAGGAATCT	1789
DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	710	CCTAACACTTGTCTTCCAGACACGTGATGCTGATAACATCTTGTATTCCAGCAA	1790	CACTTTGTGAGTGTCTGCTCCTCAATCTCAATCTCGTCTCTCTCGTCTCTCTGAGAAAG	1849
DB	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGln	581	HisPheValSerAlaArgProGlnSerGlnSerProSerProGluLysGlu	600
QY	770	GGCAGACCGCTGACCTGACCAATGCGAATGCAATACAGAAAGCTTAACTTTCGAG	1850	TCCTCTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTCCTCTTCA	1909
DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	601	SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer	620
QY	830	GGCCATGCACTCAGAATCCCATCCGCTTTCATTTCTACATCTTGAACCGCCATGACAC	1910	ATTTTGAAGGCTTTTAAAC	1927
DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	621	IleLeuLysAlaPheAsn	626
QY	890	CAGAACTCAGAGTAGCTAAATCTCCATGCCGCTTAAACACACCCGCGCAGTTTGAGAT	RESULT 7		
DB	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	ABU52568		
QY	950	TTCTTCCCGGAGCGGAGCGACCAATCATCTACTTCAGGGCTTCAGCAGGATACG	ID	ABU52568 standard; protein; 626 AA.	
DB	301	PheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	XX	ABU52568;	
QY	1010	TTGGAGCGCGCTTCAATCGGAATTCATAGATACGAGGCTGTGTAGAGAGAAT	XX	10-MAR-2003 (first entry)	
DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn	DE	Peanut Ara h1 mutant K32A.	
QY	1070	GCAGGAGGTGACAGAGGAGAGCGGACAGCGATGGAGTACTCGAGTAGTGAGAAC	KW	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;	
DB	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	KW	muten; anaphylactic food allergen; anti-allergenic; vaccine;	
QY	1130	AATGAGGAGTAGTAGTCAAGTGTCAAGGACACGTTGAGAACTTACTAAGCAGCT	XX	Homo sapiens.	
DB	361	AsnGluGlyValIleValLysValSerLysGluHisValGluLeuThrLysHisAla	PN	WO200274250-A2.	
QY	1190	AAATCGCTCTCAAGAAAGCTCCGAGAGAGGAGATATCAACCAACCAATCAACTTG	XX	26-SEP-2002.	
DB	381	LysSerValSerLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	XX	18-MAR-2002; 2002HO-US009108.	
			XX	16-MAR-2001; 2001US-0276822P.	
			XX	18-MAR-2002; 2002US-00276822.	





QY 1550 ACAGCGAGTTGAAGGAGCGATGTCTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609  
 Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520  
 QY 1610 AACGTTCTCCCAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAACCAACCCACAGA 1669  
 Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540  
 QY 1670 ATCTTCTTCAGGTGATAAGGCAATGTGTATAGACCATAGAGCAAGCAAGGAT 1729  
 Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560  
 QY 1730 TTACCATTTCTCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1789  
 Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580  
 QY 1790 CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAAT 1849  
 Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600  
 QY 1850 TCTCTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909  
 Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer 620  
 QY 1910 ATTTTGAAGCTTTTAAC 1927  
 Db 621 IleLeuLysAlaPheAsn 626  
 RESULT 8  
 ABUS2571  
 ID ABUS2571 standard; protein; 626 AA.  
 XX AC ABUS2571;  
 XX AC  
 DT 10-MAR-2003 (first entry)  
 XX DE Peanut Ara h1 mutant R91A.  
 KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;  
 KW mutein; anaphylactic food allergen; anti-allergenic; vaccine;  
 KW wound healing.  
 XX OS Homo sapiens.  
 XX OS  
 PN WO200274250-A2.  
 XX PD 26-SEP-2002.  
 XX PF 18-MAR-2002; 2002WO-US009108.  
 XX PR 16-MAR-2001; 2001US-0276822P.  
 XX PR 18-MAR-2002; 2002US-00276822.  
 XX XX (PANA-) PANACEA PHARM.  
 XX PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;  
 PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;  
 PI Rabjohn PA, Shin DS, Stanley JS;  
 XX XX WPI; 2003-018765/01.  
 XX XX  
 PT New modified anaphylactic food allergen, useful for preventing or  
 PT treating allergic reactions associated with e.g. anaphylactic allergens.  
 PS Example 5; Page; 300pp; English.  
 XX XX  
 CC The invention relates to a modified anaphylactic food allergen has an  
 CC amino acid sequence that is substantially identical to that of natural  
 CC anaphylactic food allergen, except for a cysteine residue that has been  
 CC modified so that it cannot participate in the disulphide bond. The  
 CC modification may also comprise mutation of the IgE binding sites to  
 CC reduce allergenicity. Also included are: (1) a method of making a  
 CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding

or for causing a site specific mutation in the modified anaphylactic food allergen; (3) a transgenic plant or animal expressing the modified anaphylactic food allergen; (4) a method of treating an individual by reducing the clinical response to a natural anaphylactic food allergen; and an isolated fragment of peanut allergen Ara h 1. The modified anaphylactic food allergen is useful for preventing or treating allergic reactions associated with any natural allergen such as food, insect, rubber or preferably anaphylactic allergens. It is also useful for treating wounds in mammals such as bovine, canine, feline, caprine, ovine, porcine, murine or equine species. The present sequence is a food allergen, mutated to alter its IgE binding characteristics. Note: The present sequence is not shown in the specification but was created by the indexer using information provided in the specification

XX Sequence 626 AA;

Alignment Scores:  
 Pred. No.: 5,25e-304 Length: 626  
 Score: 3280.00 Matches: 625  
 Percent Similarity: 99.84% Conservative: 0  
 Best Local Similarity: 99.84% Mismatches: 1  
 Query Match: 92.08% Indels: 0  
 DB: 6 Gaps: 0

US-10-728-051-1 (1-2032) x ABUS2571 (1-626)

QY 50 ATGAGAGGAGGGGTTTCTCCACTGATGCTGTCTGCTAGGGATCTTGTCTCGGTTTCAGTT 109  
 Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20  
 QY 110 TCTGCAACGATGCCAAGTCATCCTTACCAGAAAGAAACAGAGAACCCCTCGGCCAG 169  
 Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40  
 QY 170 AGTGCTCCAGAGTGTCAACAGGAACCGGATGACTTCAAGCAAAAGGATCGAGTCT 229  
 Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60  
 QY 230 CGTGCCACCAAGCTCGAGTATGATCTCTGTTGTCTATGATCTCGAGACACACTGGC 289  
 Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80  
 QY 290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACACGTCGCGCCGCAACCCCGAGACTAC 349  
 Db 81 ThrThrAsnGlnArgSerProGlyGluAlaThrArgGlyArgGlnProGlyAspTyr 100  
 QY 350 GATGATGACCGCGCTCAACCCGAGAGGAGGAGGCGGATGGGACAGCTGGACCG 409  
 Db 101 AspAspAspArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120  
 QY 410 AGGAGCGTCAAGAGAGAAGACTGGAGACCAACCAAGAGAGATGGAGCGCACCAAGT 469  
 Db 121 ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgGProSer 140  
 QY 470 CATCAGACGCCAGGAAATAGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529  
 Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr 160  
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACCAACCTTCTTACTTCCGCTCAAGG 589  
 Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180  
 QY 590 CGGTTTAGCACCCGCTACCGGAACCAACCGTAGGATCCGGTCTCTCGCAGAGGTTTGAC 649  
 Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200  
 QY 650 CAAAGTCAAGCGAGTTTCAGAAATCTCCAGATCCACCGTATTGTGAGATCGAGGCCAAA 709  
 Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220  
 QY 710 CCTAACACTCTTCTTCTCCAGCAGCTGATGCTGATTAACACTCTTCTTATCCAGCAA 769  
 Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240

QY	770	GGGCAAGCCACCGTGACCGTACCAATGCAATATACAGAAAGAGCTTTAATCTTCACGAG	829	QY	1850	TCTCTCAGAAAGAGGATCAAGAGGAGGAAACCAAGAGGAGGAGGTCCACTCTTTCA	1909
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260	Db	601	SerProGluLysGluAspGlnGluGluAsnGlnGlyGlySerProLeuLeuSer	620
QY	830	GGCCATGCACTCAGATCCATCCGCTTTCATTTCTTACATCTTCAACCGCATCACAA	889	QY	1910	ATTTTGAAGCTTTTAAAC	1927
Db	261	GlyHisAlaLeuArgLysProSerGlyPheIleSerTyrlleuAsnArgHisAspAsn	280	Db	621	IleLeuLysAlaPheAsn	626
QY	890	CAGAACTCTCAGAGTACGTAATCTCCATGCGCGTTTAAACACACCCCGCGAGTTTGAAGAT	949	RESULT 9			
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300	ABUS2574	ID	ABUS2574	standard; protein; 626 AA.
QY	950	TTCTTCCCGCGAGCGGAGACCAATCATCTTCTTCTCAGGGCTTCAGCAGGATACG	1009	AC	XX	ABUS2574;	
Db	301	PhePheProAlaSerArgAspGlnSerSerTyrlleuGlnGlyPheSerArgAsnThr	320	DT	XX	10-MAR-2003	(first entry)
QY	1010	TTGGAGCGCGCTTCAATCGGAATTCATAGATACGAGGCTGCTGTAGAGAGAT	1069	XX	DE	Peanut Ara h1 mutant R499A.	
Db	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn	340	KW	XX	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;	
QY	1070	GCAGGAGGTGAGCAAGAGAGAGCGGAGCGGATGGAGTACTCGGAGTGTAGTGAAC	1129	KW	XX	mucin; anaphylactic food allergen; anti-allergenic; vaccine;	
Db	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360	KW	XX	wound healing.	
QY	1130	AATGAGGAGTGTAGTCAAGTGTCAAGGAGCAGCTTGAAGACTTACTAAGCAGCT	1189	OS	XX	Homo sapiens.	
Db	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380	XX	XX	WO200274250-A2.	
QY	1190	AAATCCCTCTCAAGAAGGCTCCGAGAAGAGGAGATATCAACCAACCAATCAACTTG	1249	XX	XX	26-SEP-2002.	
Db	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400	XX	XX	18-MAR-2002; 2002WO-US009108.	
QY	1250	AGAGAAGCGAGCCGATCTTTCTAACAATTTGGGAGTATTTCAGGTGAGCCAGAC	1309	XX	XX	16-MAR-2001; 2001US-0276822P.	
Db	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420	PI	XX	18-MAR-2002; 2002US-00276822.	
QY	1310	AAGAAGAACCCCGCTTCCAGACTCGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369	PI	XX	(PANA-) PANACEA PHARM.	
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440	PI	XX	Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;	
QY	1370	CGAGCTTTGATGCTCCCACTTCAACTCAAGCCATGTTATCGTCGTCGTCACAAA	1429	PI	XX	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;	
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460	PI	XX	RabJohn PA, Shin DS, Stanley JS;	
QY	1430	GGAAGTGAACCTGAACTGCTGCTGCTAAGAAAGAGCAACAAAGAGGAGCGCGG	1489	DR	XX	WPI; 2003-018765/01.	
Db	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgArg	480	PT	XX	New modified anaphylactic food allergen, useful for preventing or	
QY	1490	GAAGAAGAGGAGGAGCAAGAGAGAGAGGAGGAGTAACACAGAGGTGCGTAGTAC	1549	PT	XX	treating allergic reactions associated with e.g. anaphylactic allergens.	
Db	481	GluGluGluGluAspGluAspGluGluGluGlySerAsnArgValArgArgTyr	500	XX	PS	Example 5; Page; 300pp; English.	
QY	1550	ACAGCGAGTTGAAGAGGAGTGTTCATCATGCCAGAGCTCATCCAGTAGGCCATC	1609	XX	XX	The invention relates to a modified anaphylactic food allergen has an	
Db	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520	CC	CC	amino acid sequence that is substantially identical to that of natural	
QY	1610	RAAGCTTCTCGAACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1669	CC	CC	anaphylactic food allergen, except for a cysteine residue that has been	
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg	540	CC	CC	modified so that it cannot participate in the disulphide bond. The	
QY	1670	ATCTCTCTTTCAGGATGAAGACAAATGTATAGACCAAGATAGACCAAGCAGGAT	1729	CC	CC	modification may also comprise mutation of the IgE binding sites to	
Db	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysasp	560	CC	CC	reduce allergenicity. Also included are: (1) a method of making a	
QY	1730	TTAGCATTCCTCGGTCGAGTCAAGTGTGAGAGCTCATCAAAACCAAGAAATCT	1789	CC	CC	modified anaphylactic food allergen; (2) a nucleotide molecule encoding	
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580	CC	CC	or for causing a site specific mutation in the modified anaphylactic food	
QY	1790	CACCTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCTCTCTCTCTCTCTCTCTCTCT	1849	CC	CC	allergen; (3) a transgenic plant or animal expressing the modified by	
Db	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600	CC	CC	anaphylactic food allergen; (4) a method of treating an individual by	

Sequence 626 AA;  
Alignment Scores: 5.25e-304 Length: 626  
Pred. No.: 3280.00 Matches: 625  
Score: 99.84% Conservative: 0  
Percent Similarity:

Best Local Similarity: 99.84%		Mismatches: 1
Query Match:	92.08%	Indels: 0
DB:	6	Gaps: 0
US-10-728-051-1 (1-2032) x ABUS2574 (1-626)		
QY	50	ATCAGAGGAGGGTTTCTCCAGTATGCTGTCTAGGATCCTTGTCTGGCTTCAGTT 109
DB	1	MetArgIYArgValSerProLeuMetLeuLeuGlyLeuValLeuAlaSerVal 20
QY	110	TCGCAACGATCCCAAGTCATCACTTACAGAGAAACACAGAACCCCTCGGCCAG 169
DB	21	SerAlaThrHisAlaIysSerProTyrGlnIysThrGluAsnProCysAlaGln 40
QY	170	AGTGCTCCAGATTTGTCAACAGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
DB	41	ArgCysLeuGlnSerCysGlnGlnProAspAspLeuIysGlnIysAlaCysGluSer 60
QY	230	CGCTGCACCAAGTCAGTATGATCTCTGTGTCTATGATCTCTCGAGGACACACTGGC 289
DB	61	ArgCysThrIysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY	290	ACCACCAACCAAGTTCCTCCAGGGAGCGACACGTCGGCCGCAACCCGAGACTAC 349
DB	81	ThrThrAsnGlnArgSerProProGlyIuArgThrArgGlyArgGlnProGlyAspTyr 100
QY	350	GATGATGACCGCGTCAACCCCAAGAGAGAGAGCGGATGGGGACCGAGCTGGACCG 409
DB	101	AspAspArgArgGlnProArgGluGlyArgGlyGlyProAlaGlyPro 120
QY	410	AGGAGCGTGAAGAGAGAAGACTGGAGACAACTGAAGAAATTTGGAGGCGCAAGT 469
DB	121	ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140
QY	470	CATCAGCAGCCAGAAATAAGCCGAGGAGAGAGAGAGAGAAACCAAGAGTGGGAACA 529
DB	141	HisGlnGlnProArgIysIleArgProGluGlyArgGluGlnGluTrpGlyThr 160
QY	530	CCAGGTAGCATGTGAGGAGAAACATCTCGAAACAACCTTTCTACTTCCGCTCAAGG 589
DB	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY	590	CGGTTTAGCACCCGCTACGGGAACCAAAACGTTAGGATCCGGTCTCTGAGAGTTTGAC 649
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY	650	CRAAGCTCAAGCAGTTTCAGATCTCCAGAAATCCAGTATTCGTCAGATCGAGGCCAA 709
DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY	710	CCTAACACTTGTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATCCAGCAA 769
DB	221	ProAsnThrLeuValLeuProIysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY	770	GGCAAGCCACCTGACCGTAGCAATGCGCAATAACAGAAAGCTTTAATCTTGACGAG 829
DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnArgIysSerPheAsnLeuAspGlu 260
QY	830	GGCATGCACTCAGATCCCATCCGTTTCAATCTTCTACATCTTGAACCGCCATGACAC 889
DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY	890	CAGAACCTCAGATGAGTAAATCTCCATCCCGTAAACACACCCCGCCAGTTTGAAGAT 949
DB	281	GlnAsnLeuArgValAlaIysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY	950	TTCTTCCCGCGAGCGAGCCAGACCAATCATCTTCTAGGAGGCTTCAGCAGGAATACG 1009
DB	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY	1010	TTGAGGCCCGCTTCAATCGGAATTCATAGATACGAGGCTGTGTAGAGAGAAAT 1069
DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY	1070	GCAGAGGTGACAAAGAGAGAGAGGCGAGCGCATGAGTACTCGAGTACTGTAGAAC 1129
DB	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
QY	1130	AATGAGGAGTCAATAGTCAAACTGTCAAAGGAGGACGTTTGAAGAACTTACTAGCAGCT 1189
DB	361	AsnGluGlyValIleValIysValSerIysGluHisValGluGluLeuThrIysHisAla 380
QY	1190	AAATCCGTCTCAAGAAAGGCTCCGAAAGAGGAGAGATATCACCAACCAATCAACTTG 1249
DB	381	LysSerValSerIysIysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY	1250	AGAGAAGCGAGCCCGATCTTTCTAACTTTGGGAAGTTATTTGAGTGAAGCCAGAC 1309
DB	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValIysProAsp 420
QY	1310	AGAGAAGCCCGAGCTTCAGGACCTGGACATGCTCACCTGTCTAGAGATCAAGAA 1369
DB	421	LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleIysGlu 440
QY	1370	CGAGCTTTTCATGCTCCACACTTCAACTCAAGGCCATGTTATCTCGTCGTCAACAA 1429
DB	441	GlyAlaLeuMetLeuProHisPheAsnSerIysAlaMetValIleValValAsnLys 460
QY	1430	GGAATCGAAACCTTGAACTCGTGTGTGAAGAAAGACAAACAGAGGGGACCGCGG 1489
DB	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY	1490	GAAGAAGAGGAGGAGCAAGACGAGAGGAGGAGGAGTAAACAGAGAGGTCGTAGTAC 1549
DB	481	GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgAlaTyr 500
QY	1550	ACAGCAGGTTCGAAGAGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
DB	501	ThrAlaArgLeuIysGlyAspValPheIleMetProAlaHisProValAlaIle 520
QY	1610	AACGCTTCTCCGAACTCCATCTGCTTGGCTTCGTTATCAACGCTCAAAACCAACACAGA 1669
DB	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
QY	1670	ATCTTCTTCCAGGTGATAGGACATCTGATAGACAGATAGAGAGCAAGCGAAGGAT 1729
DB	541	IlePheLeuAlaGlyAspIysAspAsnValIleAspGlnIleGluIysGlnAlaLysAsp 560
QY	1730	TTAGCATTCCTGGGTCCGGTGAAACAAAGTTGAGAAGCTCATCAAAACAGAGGAATCT 1789
DB	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnIysGluSer 580
QY	1790	CACCTTTGTGAGTCTCGTCTCTCAATCTCAATCTCCGTCGTCCTCTGAGAAGAG 1849
DB	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY	1850	TCTCTGAGAAAGAGATCAAGAGGAGGAAACCAAGAGGAGGAGGTCCTCCTTCA 1909
DB	601	SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY	1910	ATTTTGAAGCTTTTAAAC 1927
DB	621	IleLeuLysAlaPheAsn 626
RESULT 10		
ABUS2573		
ID	ABUS2573	standard; protein; 626 AA.
XX	ABUS2573;	
DT	10-MAR-2003	(first entry)
XX	Peanut Ara h1 mutant R109A.	
XX	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;	
KW	mucin; anaphylactic food allergen; anti-allergenic; vaccine;	

wound healing.

Homo sapiens.

WO200274250-A2.

26-SEP-2002.

18-MAR-2002; 2002WO-US009108.

16-MAR-2001; 2001US-0276822P.

18-MAR-2002; 2002US-00276822.

(PANA-) PANACEA PHARM.

Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G, Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ, Rabjohn PA, Shin DS, Stanley JS;

WPI; 2003-018765/01.

New modified anaphylactic food allergen, useful for preventing or treating allergic reactions associated with e.g. anaphylactic allergens.

Example 5; Page; 300pp; English.

The invention relates to a modified anaphylactic food allergen has an amino acid sequence that is substantially identical to that of natural anaphylactic food allergen, except for a cysteine residue that has been modified so that it cannot participate in the disulphide bond. The modification may also comprise mutation of the IGE binding sites to reduce allergenicity. Also included are: (1) a method of making a modified anaphylactic food allergen; (2) a nucleotide molecule encoding or for causing a site specific mutation in the modified anaphylactic food allergen; (3) a transgenic plant or animal expressing the modified anaphylactic food allergen; (4) a method of treating an individual by reducing the clinical response to a natural anaphylactic food allergen; and an isolated fragment of peanut allergen Ara h 1. The modified anaphylactic food allergen is useful for preventing or treating allergic reactions associated with any natural allergen such as food, insect, rubber or preferably anaphylactic allergens. It is also useful for treating wounds in mammals such as bovine, canine, feline, caprine, ovine, porcine, murine or equine species. The present sequence is a food allergen, mutated to alter its IGE binding characteristics. Note: The present sequence is not shown in the specification but was created by the indexer using information provided in the specification

Sequence 626 AA;

Alignment Scores:

Pred. No.:	5.25e-304	Length:	626
Score:	3280.00	Matches:	625
Percent Similarity:	99.84%	Conservative:	0
Best Local Similarity:	99.84%	Mismatches:	1
Query Match:	92.08%	Indels:	0
DB:	6	Gaps:	0

US-10-728-051-1 (1-2032) x ABUS2573 (1-626)

QY 50 ATGAGGGGGGGTTTCTCCACTGATGCTTGTCTAGGATCCTTCTCGCTTCAGTT 109

DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20

QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAAGAAAAACAGAAACCCCTGCGCCAG 169

DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

QY 170 AGTGCTCCAGATTGTCAACAGAACCCGATGACTTGAAGCAAAAGCATGCGAGTCT 229

DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60

QY 230 CGCTGCACCAAGCTCAGATGATCTCTGTTGTCTATGATCCTCGAGGACACATGGC 289

Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGACACAGTGGCGGCCCAACCCGAGACTAC 349

DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100

QY 350 GATGATGACCCGCGTCAACCCCGAAGAGAGAGAGGCGCGATGGGACCACTGGACCG 409

DB 101 AspAspArgArgGlnProArgAlaGluGluGlyGlyArgTrpGlyProAlaGlyPro 120

QY 410 AGGGAGCGTGAAGAGAGAAAGACACTGGAGACAAACCAAGAGAAGATTGGAGGCGCAACT 469

DB 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer 140

QY 470 CATCAGACCCACGGAATAAAGCCCGAAGAGAGAGAGAGAAACAAAGATGGGGAACA 529

DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnTrpGlyThr 160

QY 530 CCAGGTAGCCATGTGAGGGAAGAAAACATCTCGAAACAACCCCTTCTACTTCCCGTCAAGG 589

DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180

QY 590 CGGTTTAGCACCCGCTACGGGAACCAACACGTTAGATCCGGGTCTCGAGAGTTGAC 649

DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200

QY 650 CAAAGTCAAGCAGTTTTCAGAAATCTCAGAAATCCACGTAATGTGTGAGATCCAGGCAAA 709

DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220

QY 710 CTTAACACTCTTGTCTTCCCAAGCAGCTGATGCTGATAACATCTTCTTCTTATCCAGCAA 769

DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240

QY 770 GGGCAGCCACCGTCAGCTAGCAATGCGCAATACAGAAAGCGTTTAAATCTTGACGAG 829

DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260

QY 830 GGCCATGCACTCAGAAATCCATCCCGTTCATTTCTTCTACATCTTGAACCGCATCACAAC 889

DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280

QY 890 CAGAACTCAGAGTAGCTTAAATCTCCATGCGCCGTTTAAACACCCGCGCAGTTTGAAGAT 949

DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300

QY 950 TTTCTTCCGCGCAGCAGCGACCAATCATCTTCTGACGGCTTTCAGGAGGATACG 1009

DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320

QY 1010 TTGGAGGCGCCCTTCAATGCGGAATTCATATGAGATACGAGGGTCTGTGTAGAAGAAAT 1069

DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340

QY 1070 GCAGGAGGTGACCAAGAGAGAGAGCGGACGAGCGATGGAGTACTCGGAGTAGTGAGAAC 1129

DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgTrpSerThrArgSerSerGluAsn 360

QY 1130 AATGAGGAGTCATAGTCAAACTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCACGCT 1189

DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380

QY 1190 AAATCCGCTCTCAAAGAAAGCGTCCGAAGAGAGGAGGATATCACCACCAATCACTTG 1249

DB 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400

QY 1250 AGAAGAGCGAGCCCGATCTTTCTAACTTTGGGAAGTATTATTCAGGTGAAGCCAGAC 1309

DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420

QY 1310 AAGAGAAGACCCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTGAGATCAAGAA 1369

DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLeuLysGlu 440

QY 1370 GGAGCTTTGATGCTCCACACTTCACTCAACTCAAGAGCCCATGTTATCGTCGTCGCAACAAA 1429  
|||  
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460  
QY 1430 GGAACGGAACCTTGAACTCGCTGCTGTAAAGAGAGCAACACAGAGGGGACGCGG 1489  
|||  
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480  
QY 1490 GAAGAAGAGGAGGACGAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549  
|||  
Db 481 Glu 500  
QY 1550 ACAGCGAGGTGAAGAGCGGATGTTTCATCATGTCAGCAGCTCATCCAGTAGCCATC 1609  
|||  
Db 501 ThrAlaArgLeuLysGluGlyArgValPheIleMetProAlaAlaHisProValAlaIle 520  
QY 1610 AACGCTTCCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCCACAGA 1669  
|||  
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540  
QY 1670 ATCTTCCTTCAGGTGATAAGGACAATGTGATAGACCAAGATAGAGAGCAAGCGAAGGAT 1729  
|||  
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560  
QY 1730 TTAGCATTCCTCGGTGGAACAAAGTTGAGAAGCTCATCAAAACCAAGAGGAATCT 1789  
|||  
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580  
QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTCAAGAGAG 1849  
|||  
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600  
QY 1850 TCCTCTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1909  
|||  
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyGlyGlyGlyGlyGlyGly 620  
QY 1910 ATTTTGAAGCGTTTAAAC 1927  
|||  
Db 621 IleLeuLysAlaPheAsn 626

RESULT 11  
ADM12053  
ID ADM12053 standard; protein; 626 AA.  
XX AC ADM12053;  
XX DT 20-MAY-2004 (first entry)  
XX DE Arachis hypogaea 1 (Ara hi) protein.  
XX KW antigen presentation enhancing hybrid polypeptide; mammalian ii-key;  
KW MHC class ii; antibacterial; virucide; fungicide; antirheumatic;  
KW antithratic; neuroprotective; dermatological; immunosuppressive;  
KW antiinflammatory; antidiabetic; antithyroid; immune;  
KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;  
KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;  
KW scleroderma; dermatomyositis; pemphigus.  
XX OS Arachis hypogaea.  
XX PN US2003235594-A1.  
XX XX  
XX PD 25-DEC-2003.  
XX XX  
XX PF 17-SEP-2002; 2002US-00245871.  
XX XX  
XX PR 14-SEP-1999; 99US-00396813.  
XX PR 17-JUL-2002; 2002US-00197000.  
XX XX  
XX PA (ANTI-) ANTIGEN EXPRESS INC.  
XX XX  
XX PI Humphreys R, Xu M;

XX WPI; 2004-070554/07.

XX Novel ii-key/antigen presentation enhancing hybrid polypeptide, useful  
PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus  
PT erythematosus and diabetes mellitus.

XX Example 1; Page 16-17; 87pp; English.

XX The invention relates to a novel antigen presentation enhancing hybrid  
CC polypeptide. The novel polypeptide has an N-terminal element consisting  
CC of 4-16 residues of a mammalian ii-key peptide and its non-N-terminal  
CC deletion modifications, a chemical structure covalently linking the N-  
CC terminal element to an MHC class ii-presented epitope of a C-terminal  
CC element. The C-terminal element comprises an antigenic epitope, which  
CC binds to an antigenic peptide binding site of an MHC class II molecule.  
CC The antigen presentation enhancing hybrid polypeptide has the following  
CC activities: antibacterial, virucide, fungicide, antirheumatic,  
CC antithratic, neuroprotective, dermatological, immunosuppressive,  
CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation  
CC enhancing hybrid polypeptide is useful for modulating the immune response  
CC in an individual and for treating infections (such as bacteria, virus,  
CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus  
CC erythematosus, diabetes mellitus, myasthenia gravis, autoimmune  
CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence  
CC represents a mammalian ii key related protein of the invention.

XX SQ Sequence 626 AA;

Alignment Scores:  
Pred. No.: 6,54e-304 Length: 626  
Score: 3279.00 Matches: 625  
Percent Similarity: 99.84% Conservative: 0  
Best Local Similarity: 99.84% Mismatches: 1  
Query Match: 92.06% Indels: 0  
DB: 8 Gaps: 0

US-10-728-051-1 (1-2032) x ADM12053 (1-626)

QY 50 ATGAGAGGAGGGGTTTCTCCACTGATGCTGTTGCTAGGGATCCTTGCTCGGTTCAAGTT 109  
|||  
Db 1 MetArgLysValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20  
QY 110 TCTGCAACGCATCCCAAGTCATCACCTTACCAGAGAGAAACAGAGAACCCCTCGCCACG 169  
|||  
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40  
QY 170 AGTGCTCTCCAGAGTTGTCACAGGACCGGATGACTTGAGCAAAAGGATCGGAGTCT 229  
|||  
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlyLysAlaCysGluSer 60  
QY 230 CGCTGCACCAAGCTCCAGTATGATCCTCGTTGTCTATGATCCTCGAGGACACACTGGC 289  
|||  
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80  
QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGCGGACACGTCGCGCCCAACCCGAGACTAC 349  
|||  
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100  
QY 350 GATGATCAGCCCGCTCAACCCCGAAGAGAGAGGAGGCGGATCGGAGGACGAGTGGACCG 409  
|||  
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrProAlaGlyPro 120  
QY 410 AGGGAGCGTGAAAG 469  
|||  
Db 121 ArgGluArgGluArgGluGluAspTyrPargGlnProArgGluAspTyrPargArgProSer 140  
QY 470 CATCAGCAGCCCGAAAAATAAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529  
|||  
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr 160  
QY 530 CGAGGTAGCCATGTGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 589  
|||

161	ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
590	CGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGATCCGGGCTCTGCAGAGATTGAC	649
181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
650	CAAAGGTCGAAGCAGTTTCAGAAATCTCAGAAATCAACCGTATTGTGCGAGATCGAGGCCAAA	709
201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
710	CCTAAACACTCTTGTTCTCCCAAGCAGCGCTGATGCTGATAATCATCTCTGTATTATCCAGCAA	769
221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
770	GGGCAAGCCACCGTGACCGTAGCAATAACAGAAAGAGCTTTAAATCTTTGACGAG	829
241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
830	GGCCATGCATCAGATCCCATCCCGTTTCATTCTCATCATCTTCAACCGCATGACAAAC	889
261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
890	CAGAACTCTCAGAGTAGCTAAATCTCCATCCCGCTTAAACACCCCGGCAGTTTGAGAT	949
281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
950	TTCTTCCGGCAGCAGCCGAGACCAATCATCTACTCTCAGGGCTTCAGCAGGAATACG	1009
301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
1010	TTGAGCGCGCTTCAATGCGGAATCAATGAGATACGGAGGCTGCTGTAGAAGAGAT	1069
321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn	340
1070	GCAGGAGTGACGAAGAGGAGAGGCGCAGTGCAGTACTCCGAGTAGTGAGAAC	1129
341	AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerGluAsn	360
1130	AATGAAGAGTGATAGTCAAGTGTCAAGCGAGCAGCTTGAAGAACTTACTAAGCACGCT	1189
361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
1190	AAATCCGCTTCAAAAGAAAGGCTCCGAAGAAAGAGGAGATATACCAACCCATCACTTG	1249
381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
1250	AGAGAAGCGGACCCGATCTTCTTAAACAATTTGGGAAGTATTTCAGGTGAAGCCAGAC	1309
401	ArgGluGlyGluProAspLysSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
1310	AAGAAAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAGAA	1369
421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
1370	GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCGTCGTCAACAAA	1429
441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
1430	GGAACTGGAAACCTTGAACCTCTGTGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGG	1489
461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgIleArgArg	480
1490	GAAGAAGAGGAGGAGCAAGACGAGAGGAGGGAAGTACACAGAGGTGCGTAGGTAC	1549
481	GluGluGluGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
1550	ACAGCGAGGTGAAGGAAGCGCATGTTGTTCATATGCCAGAGCTCATCCAGTAGCCATC	1609
501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
1610	AACGCTTCTCCGAATCCATCTGTGCTGGCTATCAGCTGTGAAGAACCAACAGAGA	1669
521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg	540

Qy	1670	ATCTCTCTTCGAGTGCATTAAGGACAATGTGATAGACCATAGAGAGCAAGCGRAGGAT	1720
Db	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluysGlnAlaLysAsp	560
Qy	1730	TTAGCAATTCCTCGGTGGTGAACAAGTGTGAGAAGTCATCAAAAACCCAGAAAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
Qy	1790	CACCTTTGTGAGTCTCGTCTCTCAATCTCAATCTCCGTCCTCTCTGGAAGAG	1849
Db	581	HisPheValSerAlaArgProGlnSerGlnSerProSerProGlnLysGlu	600
Qy	1850	TCTCTCTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTCCTCACTCTTCA	1909
Db	601	SerProGlnLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
Qy	1910	ATTTTGAAGGCTTTTAAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
RESULT 12			
ABU52569			
ID	ABU52569	standard; protein; 626 AA.	
XX	AC	ABU52569;	
XX	DT	10-MAR-2003 (first entry)	
XX	DE	Peanut Ara h1 mutant D52A.	
XX	KW	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;	
XX	KW	muten; anaphylactic food allergen; anti-allergenic; vaccine;	
XX	KW	wound healing.	
XX	OS	Homo sapiens.	
XX	PN	WO200274250-A2.	
XX	PD	26-SEP-2002.	
XX	PF	18-MAR-2002; 2002WO-US009108.	
XX	PR	16-MAR-2001; 2001US-0276822P.	
XX	PR	18-MAR-2002; 2002US-00276822.	
XX	PA	(PANA-) PANACEA PHARM.	
XX	PA	Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;	
XX	PI	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;	
XX	PI	Rabjohn PA, Shin DS, Stanley JS;	
XX	XX	WPI; 2003-018765/01.	
XX	XX	New modified anaphylactic food allergen, useful for preventing or	
XX	XX	treating allergic reactions associated with e.g. anaphylactic allergens.	
XX	XX	Example 5; Page; 300pp; English.	
XX	XX	The invention relates to a modified anaphylactic food allergen has an	
XX	XX	amino acid sequence that is substantially identical to that of natural	
XX	XX	anaphylactic food allergen, except for a cysteine residue that has been	
XX	XX	modified so that it cannot participate in the disulphide bond. The	
XX	XX	modification may also comprise mutation of the IgE binding sites to	
XX	XX	reduce allergenicity. Also included are: (1) a method of making a	
XX	XX	modified anaphylactic food allergen; (2) a nucleotide molecule encoding	
XX	XX	or for causing a site specific mutation in the modified anaphylactic food	
XX	XX	allergen; (3) a transgenic plant or animal expressing the modified	
XX	XX	anaphylactic food allergen; (4) a method of treating an individual by	
XX	XX	reducing the clinical response to a natural anaphylactic food allergen;	
XX	XX	and an isolated fragment of peanut allergen Ara h 1. The modified	
XX	XX	anaphylactic food allergen is useful for preventing or treating allergic	
XX	XX	reactions associated with any natural allergen such as food, insect,	



CC rubber or preferably anaphylactic allergens. It is also useful for  
 CC treating wounds in mammals such as bovine, canine, feline, caprine,  
 CC ovine, porcine, murine or equine species. The present sequence is a food  
 CC allergen, mutated to alter its IgE binding characteristics. Note: The  
 CC present sequence is not shown in the specification but was created by the  
 CC indexer using information provided in the specification

XX  
 SQ Sequence 626 AA;

# Alignment Scores:

Pred. No.: 8.15e-304 Length: 626  
 Score: 3278.00 Matches: 625  
 Percent Similarity: 99.84% Conservative: 0  
 Best Local Similarity: 99.84% Mismatches: 1  
 Query Match: 92.03% Indels: 0  
 DB: 6 Gaps: 0

US-10-728-051-1 (1-2032) x ABUS2569 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCATGATGCTGCTGCTAGGATCTCTGTGCTGGCTTCAGTT 109  
 DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20  
 QY 110 TCTGCAACGATGCCAGTTCATCACCTTACAGAGAAACAGAGAACCCCTCGGCCAG 169  
 DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40  
 QY 170 AGGTGCTCCAGAGTTGTCACAGGAACCGGATGACTTTGAAGCAAAAGGATCGAGTCT 229  
 DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAlaLeuLysGlnLysAlaCysGluSer 60  
 QY 230 CGCTGCAACCAAGTTCGAGTATGATCTCTGTTGTCTATGATCTTCAGGACACATGCG 289  
 DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80  
 QY 290 ACACCAACCAAGTTCCTCCAGGGAGCGGACGTCGCGCCGCAACCCGAGACTAC 349  
 DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100  
 QY 350 GATGATACCCGCTCAACCCGAAAGAGAGAGAGCGGCGATGGGACACGATGGACCG 409  
 DB 101 AspAspAspArgGlnProArgGluGluGlyGlyArgTyrProAlaGlyPro 120  
 QY 410 AGGAGCGTGAAGAGAAAGACTGGAGACAACCAAGAGAAGATTGGAGCGCACCAAGT 469  
 DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgGlnProSer 140  
 QY 470 CATCAGACGACGGAATATAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529  
 DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTyrPglyThr 160  
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACCAACCTTTCTACTTCCGCTCAAG 589  
 DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180  
 QY 590 CGGTTTAGCCCCCTACGGAAACCAAAACGGTAGGATCCGGGTCTTCGAGAGGTTTAC 649  
 DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200  
 QY 650 CAAAGGTCAAGGAGTTTCAGATCTCCAGATCCAGATCCAGATCCAGATCCAGATCC 709  
 DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220  
 QY 710 CCTAACACTCTTGTCTTCCCAAGCAGCTGATCTGATTAACATCTTGTATTCCAGCAA 769  
 DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240  
 QY 770 GGGCAAGCCCGTAGCCGTAAGCAAAATGGCAATTAACAGAAAGACTTTAATCTTTGACG 829  
 DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260  
 QY 830 GGCCATGCATCAGAAATCCCATCCGGTTTCAATCTTCTACATCTTGAACCCGCTACAC 889

DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280  
 QY 890 CAGAACTCAGAGTAGCTAAATCTCCATCCCGTTTAAACACACCCGCGCAGTTTGAGGAT 949  
 DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300  
 QY 950 TTCTTCCCGGAGCAGCCGAGACCAATCATCTACTTGCAGGGCTTCACAGGATACG 1009  
 DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320  
 QY 1010 TTGGAGCGCGCTTCAATCGGAATTCATAGATACGAGGAGTCTGTTAGAGAGAAT 1069  
 DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340  
 QY 1070 GCAGAGGTGAGCAAGAGGAGAGAGCGGAGAGCGATGGAGTACTCGGAGTAGTGAAC 1129  
 DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgTyrSerThrArgSerSerGluAsn 360  
 QY 1130 AATGAAGGAGTGTAGTCAAGTGTCAAGGAGCAGGTTGAAGAACTTACTAGCACGCT 1189  
 DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380  
 QY 1190 AATCCGCTCTCAAAGAAAGGCTCCGAAGAGAGAGAGATATCACCAACCCCAATCACTTG 1249  
 DB 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400  
 QY 1250 AGAAGGCGAGCCGATCTTCTTAAACAATTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309  
 DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420  
 QY 1310 AAGAAAGACCCGAGCTTCAGGACCTGGACATGCTACCTGCTGTGTAGAGATCAAGAA 1369  
 DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440  
 QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGCCATGTTATCGTCTGCTCAACAAA 1429  
 DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460  
 QY 1430 GGAATCGAAACCTTCAACTCTGGCTGTAAAGAAAAGACAACACAGAGGGAGCGCGG 1489  
 DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 480  
 QY 1490 GAAGAGAGGAGACCAAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549  
 DB 481 GluGluGluGluAspGluGluGluGluGluGlySerAsnArgGluValArgTyr 500  
 QY 1550 ACAGCAGGTTGAAGAGAGCGGATGTTTCATCATGCGCAGCTCATCCAGTAGCCATC 1609  
 DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisProValAlaIle 520  
 QY 1610 AACGCTTCTCCGAATCCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACACACAGA 1669  
 DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540  
 QY 1670 ATCTTCTTCAGGCTAAGGACATGTGATGACAGATAGAGAGAGAGAGAGAGAGAGAT 1729  
 DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560  
 QY 1730 TTAGCATTTCCCTCGGTGAACAAAGTTGAGAAGCTCATCAAAACACAGAGAGAAATCT 1789  
 DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580  
 QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGCTGCTCTCTGAGAAAGAG 1849  
 DB 581 HisPheValSerAlaArgProGlnSerGlnSerProSerProSerProGluLysGlu 600  
 QY 1850 TCTCTCGAGAAAGAGATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGAGGAGGAGG 1909  
 DB 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620  
 QY 1910 ATTTTGAAGGCTTTTAAAC 1927  
 DB 621 IleLeuLysAlaPheAsn 626

## RESULT 13

ABU52572  
ID ABU52572 standard; protein; 626 AA.

XX AC

XX ABU52572;

XX XX

DT 10-MAR-2003 (first entry)

XX DE

XX Peanut Ara h1 mutant D103A.

XX XX

XX Allergy; Ara h1; Ara h2; Ara h3; IGE binding site; peanut; mutant;

KW KM

XX mutein; anaphylactic food allergen; antiallergenic; vaccine;

XX KW

XX wound healing.

XX XX

XX OS

XX Homo sapiens.

XX XX

XX PN

XX W0200274250-A2.

XX XX

XX PD

XX 26-SEP-2002.

XX XX

XX PF

XX 18-MAR-2002; 2002WO-US009108.

XX XX

XX PR

XX 16-MAR-2001; 2001US-0276822P.

XX PR

XX 18-MAR-2002; 2002US-00276822.

XX XX

XX PA

XX (PANA-) PANACEA PHARM.

XX XX

XX Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;

PI PI

XX Caprade CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;

PI PI

XX Rabbjohn PA, Shin DS, Stanley JS;

XX DR

XX WPI; 2003-018765/01.

XX PT

XX New modified anaphylactic food allergen, useful for preventing or

PT PT

XX treating allergic reactions associated with e.g. anaphylactic allergens.

XX XX

XX PS

XX Example 5; Page; 300pp; English.

XX CC

XX The invention relates to a modified anaphylactic food allergen has an

CC CC

XX amino acid sequence that is substantially identical to that of natural

CC CC

XX anaphylactic food allergen, except for a cysteine residue that has been

CC CC

XX modified so that it cannot participate in the disulphide bond. The

CC CC

XX modification may also comprise mutation of the IGE binding sites to

CC CC

XX reduce allergenicity. Also included are: (1) a method of making a

CC CC

XX modified anaphylactic food allergen; (2) a nucleotide molecule encoding

CC CC

XX or for causing a site specific mutation in the modified anaphylactic food

CC CC

XX allergen; (3) a transgenic plant or animal expressing the modified

CC CC

XX anaphylactic food allergen; (4) a method of treating an individual by

CC CC

XX reducing the clinical response to a natural anaphylactic food allergen;

XX and an isolated fragment of peanut allergen Ara h 1. The modified  
XX anaphylactic food allergen is useful for preventing or treating allergic  
XX reactions associated with any natural allergen such as food, insect,  
XX rubber or preferably anaphylactic allergens. It is also useful for  
XX treating wounds in mammals such as bovine, canine, feline, caprine,  
XX ovine, porcine, murine or equine species. The present sequence is a food  
XX allergen, mutated to alter its IGE binding characteristics. Note: The  
XX present sequence is not shown in the specification but was created by the  
XX indexer using information provided in the specification

XX SQ

XX Sequence 626 AA;

XX

XX Alignment Scores:

XX Pred. No.:

XX Score:

XX Percent Similarity:

XX Best Local Similarity:

XX Query Match:

XX DBs:

XX

XX

XX

XX

XX

XX

XX

XX

	8,15e-304	Length:	626
Pred. No.:	3278,00	Matches:	625
Score:	99.84%	Conservative:	0
Percent Similarity:	99.84%	Mismatches:	1
Best Local Similarity:	92.03%	Indels:	0
Query Match:	92.03%	Gaps:	0
DBs:	6		

US-10-728-051-1 (1-2032) x ABU52572 (1-626)

50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTAGGATCCTTGCTGGCTTCAGTT 109

QY

	Db
1	MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
110	TCTGCAACGCGATGCCAAGTATCATACCTTACCAGAGAGAAACAGAGAAACCTCGGCCAG 169
21	SerAlaThrHisAlaLysSerProTyrGlnLysThrGluAsnProCysAlaGln 40
170	AGGTGCTCTCCAGAGTTGTCAACAGAGAACCGGATGACTTGAAGCAAAAGGCGATGCGAGTCT 229
41	ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
230	CGTGCACCAAGCTCGAGTATGATCTCGTGTGCTCTATGATCTCTCGAGGACACACTGGC 289
61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
290	ACCACCAACCAACGTTCCCTCCAGGAGGAGGACGTCGCGCCCAACCCGAGAGATAC 349
81	ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
350	GATGATGACCCGCTCAACCCCGAAGAGAGAGAGAGCCGATGGGACCAAGCTGGACCG 409
101	AspAspAlaArgArgGlnProArgArgGluGlyArgTyrGlyProAlaGlyPro 120
410	AGGAGCGTGAAG 469
121	ArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGlu 140
470	CATCAGCAGCCAGCAAAATAAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr 160
530	CCAGGTAGCATGTGAGGAG 589
161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
590	CGGTTAGCACCCTGACGGAG 649
181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
650	CAAAGTCAAGGACGTTTCAGAAATCCAGAAATCCAGAAATCCAGAAATCCAGAAATCCAG 709
201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
710	CTAACACTCTTGTCTTCCAGAGCAGCTGATGATGATGATGATGATGATGATGATGATGAT 769
221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
770	GGGCAAGCCACGTCAGCTAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 829
241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
830	GGCCATGCACTCAGAAATCCCATCCGTTTTCATTTCTTCTACATTTCTTCTTCTTCTTCT 889
261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
890	CAGAACTCAGAGTAGCTAAATCTCCATGCGCGCTTACACACACCGCGCGCTTACAGGAT 949
281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
950	TTCTTCCCGGAG 1009
301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
1010	TTGGAGGCGCTTCAATCGGAAATTCATAGATACGAGGAGGAGGAGGAGGAGGAGGAGGAT 1069
321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
1070	GCAGGAGCTGACCAAG 1129
341	AlaGlyGlyGluGlnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 360
1130	AATGAAGGAGTGTAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGT 1189

```
Db 361 AsnGluGlyValIleValIleValIleValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGCTCTCAAGAAAGCTCCGAGAAAGAGGAGATATACCAACCAATCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGCGAGCCGATCTTTCTAACCACTTTGGGAGTTATTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCAGCTTCCAGGACCTGACATGCTACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCGTCGCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAAGTGAACCTTGAACCTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGACGGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuAlaValArgLysGluGlnGlnArgGlyArg 480
QY 1490 GAAGAAGAGGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGluGluGluGluGluValArgGly 500
QY 1550 ACAGCGAGGTGAAGGAGCGGATGTGTTTCATCATGCCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAATCCATCTGCTTGGCTTGGTATCAACCTGAAACCAACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTCTCAGGTATAGGACAAATGTGATACCCAGATACAGAACGAGGAGGAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTGGTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTGTGAGTCTGCTCTCAATCTCAATCTCAATCTCCGTCGTCCTCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCAGAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1909
Db 601 SerProGluLysGluAspGlnGlnGluGluAsnGlnGlyGlyGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 14
AAW22150
ID AAW22150 standard; protein; 626 AA.
AC AAW22150;
XX AAW22150;
DT 17-OCT-2003 (revised)
DT 29-DEC-1997 (first entry)
XX Peanut allergen Ara hi.
XX Peanut; seed storage protein; allergen; allergy; hypersensitivity;
XX vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
XX ELISA; analysis; Ara hi.
XX Arachis hypogaea; strain Flotrunner.
XX Key Location/Qualifiers
FT Peptide 1..22
```

```
FT Protein /label= Sig_peptide
FT 23..626
FT /label= Mat_protein
FT Modified-site 521..523
FT /note= "N-glycosylation site"
XX
PN WO9724139-Al.
XX
XX 10-JUL-1997.
XX
XX 23-SEP-1996; 96WO-US015222.
XX
XX 29-DEC-1995; 95US-0009455P.
XX 04-MAR-1996; 96US-00610424.
XX (UYAR-) UNIV ARKANSAS.
XX
XX Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;
XX WPI; 1997-363453/33.
XX DR N-PSDB; AAT76613.
XX
XX Peanut allergens Ara hi and Ara hII - used for vaccination and in two-
XX site monoclonal antibody based ELISA.
XX
XX Claim 31; Page 172; 354pp; English.
XX
XX This polypeptide comprises major peanut allergen Ara hi (AAW22149). Its
XX sequence was deduced from cDNA clone p41b (AAT76613), isolated from
XX peanut seed cDNA using a primer (see AAT76616) based on an isolated Ara
XX hi peptide (see AAW24206). The sequence shows significant homology with
XX the vicilin family of seed storage proteins of other legumes. The
XX allergen is recognised by serum IGE from a large proportion of
XX individuals with peanut hypersensitivity. Ara hi and Ara hII (see
XX AAW24164) can be used to raise monoclonal antibodies which are used in a
XX specific two-site Mab ELISA for the detection of Ara hi or Ara hII
XX (claimed). IGE-binding Ara hi antigen epitopes (see AAW24165-87) may be
XX used in vaccines to protect against allergic reactions to peanut
XX allergens, e.g. anaphylactic shock. (Updated on 17-OCT-2003 to
XX standardise OS field)
XX
XX Sequence 626 AA;
```

## Alignment Scores:

```
Pred. No.: 1.97e-303 Length: 626
Score: 3274.00 Matches: 624
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 2
Query Match: 91.91% Indels: 0
DB: 2 Gaps: 0
```

US-10-728-051-1 (1-2032) x AAW22150 (1-626)

```
QY 50 ATGAGAGGGAGGGTTTCTCCACTGATGCTTGTGCTAGGGATCCTTCTCGGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAAGAAAAACAGAGAACCCCTCGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGACCAAGCTCCAGTATGATCCTCGTTGTGTCTATGATCTCGAGGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGCTTCCCTCCAGGGGAGCGGACACGTCGGCCGCCCAACCCGGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
```

Qy	350	GATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGCGATGGGGACCACTGGACCG	409
Db	101	AspAspAspArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro	120
Qy	410	AGGAGCGGTGAAGAGAAGAAGACTGGAGACAAACCAAGAGAAGATTGGAGGGACCAAGT	469
Db	121	ArgGluArgGluArgGluGluAspTTrpArgGlnProArgGluAspTrpArgArgProSer	140
Qy	470	CATCACGACCCGCGAAATAAAGCCCGAAGGAAGAGAGAGAAACAAGAGTGGGGAACA	529
Db	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr	160
Qy	530	CCAGGTAGCCATGTGAGGGAAGAAAATCTCGGAACAACCCCTTTCTATCCCGTCAGG	589
Db	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPhefyrPheProSerArg	180
Qy	590	CGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCGGGTCTCGCAGAGGTTTGAC	649
Db	181	ArgPheSerThrArgfyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
Qy	650	CAAAAGGTCAAGCGAGTTTCAGAATCTCCAGATCACCGTATTGTGCAGATCGAGGCCAAA	709
Db	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
Qy	710	CCTAACTCTGTTCTTCCAGACCGCTGATGCTCATAAACATCCTGCTTATCCAGCAA	769
Db	221	ProAsnThrLeuValLeuProlysHisAlaAspAlaAspAsnIleuValIleGlnGln	240
Qy	770	GGGCAAGCCACCGTAGCCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG	829
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
Qy	830	GGCATTGCACTCAGAAATCCCATCCGHTTTCATTTCTTCTAGATCTTTGAACCGCATGACAAC	889
Db	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerfyrIleuAsnArgHisAspAsn	280
Qy	890	CAGAACTTCAGATGAGTAAATCTCAATGCCCGTTAAACAACCCCGCCAGTTTGAGGAT	949
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValThrGlnProGlyGlnPheGluAsp	300
Qy	950	TTCTTCCGGCGAGCGAGCGAGACCAATCATCTTACTTGACGGCTTCAGCAGGATACG	1009
Db	301	PhePheProAlaSerSerArgAspGlnSerfyrLeuGlnGlyPheSerArgAsnThr	320
Qy	1010	TTGAGGCGCGCTTCAATGCGGAATTCATGAGATACGGAGGTCGTGTTAGAGAGAGAT	1069
Db	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn	340
Qy	1070	GCAGAGGTGACGAAGAGAGAGAGGCGAGGCGATGGAGTACTCGAGTAGTGAGAAC	1129
Db	341	AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360
Qy	1130	AATGAAGAGGTATAGTCAAAAGTGTCAAAGGAGACGCTTGAAAGAACTTACTAAGCACGCT	1189
Db	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
Qy	1190	AAATCCGCTCTCAAGAAAGGCTCCGAAGAAGAGGAGGATATCATCAACCCATCAACTTG	1249
Db	381	LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu	400
Qy	1250	AGAGAGGCGGACCGGATCTTCTAAACAATTTGGGAGTTATTGTGAGGTGAAGCCAGAC	1309
Db	401	ArgGluGlyGluProAspLysSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
Qy	1310	AAGAAGAACCCCGACTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
Qy	1370	GGAGCTTTGATGCTCCGACACTTCAACTCAAAGGCCATGGTTATCTGTCGTCAACAAA	1429
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
Qy	1430	GGAACTGGAAACCTTTGAATCTGTGCTGTGTAAAGAAAGACCAACAAGAGCGGACGGCGG	1489

Db	461	GlyThrGlyAsnLeuGluLeuValAlaValAraGlyLysGluGlnGlnInargGlyAraG9	480
Qy	1490	GAAGAAGACGAGCAGCAAGACCAAGAAAGAGGAGGAAAGTAAACAGAGAGGTGCCTAGGTAC	1549
Db	481	GluGluGluGluAspGluGluGluGluGlySerAsnArgGluValArgAsgTyr	500
Qy	1550	ACAGCGAGGTGAAGAAAGCGAGTGTTTCATCATGCCACGACGCTCATCCAGTAGCCATC	1609
Db	501	ThrAlaAraGlyLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
Qy	1610	RACGCTTCCTCCGAACCTCATCTGCTTGCTTCGGGTATCAACGCTCAAAAACAACCACAGA	1669
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisAraG	540
Qy	1670	ATTCTTCCTTCGAGGTGATAAGGACCAATGTGATAGACAGATAGAGAACGAAAGGAT	1729
Db	541	IlePheLeuAlaGlyAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
Qy	1730	TTAGCAATTCCTGGGTCCGGTGAACAAAGTTGAGAACTCATCAAAAAACAGAAAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
Qy	1790	CACCTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCCCGTCTCTCTGAGAAAGAG	1849
Db	581	HisPheValIserAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
Qy	1850	TCTCTGAGAAAGAGATCAAGAGCGAGGAAACCAAGGAGGAGGTCCACTCTCTTCA	1909
Db	601	SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
Qy	1910	ATTTTGAAGGCTTTTAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
<hr/>			
RESULT 15			
ID	ADG27464	ADG27464 standard; protein; 625 AA.	
XX	XX	ADG27464;	
AC	ADG27464;		
DT	26-FEB-2004	(first entry)	
DE	Peanut Ara h1 cDNA clone P4lb protein.		
KW	Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycinin A2Bia; Jug n1;		
KW	antiallergic; vulnerary; anaphylactic food allergen; IgE; allergy; wound.		
OS	Arachis hypogaea.		
PN	US2003202980-A1.		
PD	30-OCT-2003.		
PF	18-MAR-2002; 2002US-00100303.		
XX	29-DEC-1995;	95US-0009455P.	
PR	23-SEP-1996;	96US-00717933.	
PR	31-JAN-1998;	98US-0073283P.	
PR	13-FEB-1998;	98US-0074590P.	
PR	13-FEB-1998;	98US-0074624P.	
PR	13-FEB-1998;	98US-0074633P.	
PR	29-JUN-1998;	98US-00106872.	
PR	27-AUG-1998;	98US-00141220.	
PR	13-NOV-1998;	98US-00191593.	
PR	29-JAN-1999;	99US-00240557.	
PR	29-JAN-1999;	99US-00241101.	
PR	11-FEB-1999;	99US-00248673.	
PR	11-FEB-1999;	99US-00248674.	
PR	02-MAR-1999;	99US-0122450P.	
PR	02-MAR-1999;	99US-0122452P.	
PR	02-MAR-1999;	99US-0122560P.	
PR	02-MAR-1999;	99US-0122565P.	



Db 361 AsnGluGlyValIleValIleValSerLysGluHieValGluLeuThrLysHisAla 380  
QY 1190 AAATCCGCTCTCAAGAAAGGCTCCGAAGAGGAGATATACCAACCCCAATCAACTTG 1249  
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400  
QY 1250 AGAGAGGCGAGCCGATCTTTCTAACTTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309  
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnAsnPheGlyLysLeuPheGluValLysProAsp 420  
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGGAATCATGCTCACCTGTGTAGATCAAGAA 1369  
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440  
QY 1370 GGAGCTTTGATGCTCCACACTTCACTCAAGCCCATGGTTATCGTCGTCACAAA 1429  
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460  
QY 1430 GGAACTGGAACCTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGCGG 1489  
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 480  
QY 1490 GAAGAAGAGGAGGACGAGACGAAAGAGAGAGAGAGTAACAGAGAGGTGCGTAGGTAC 1549  
Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500  
QY 1550 ACAGCGAGGTTGAGGAGGCGATGTGTTTCATCGCAGCAGCTCATCCAGTAGCCATC 1609  
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520  
QY 1610 AACGCTTCTCCGAACCTCCATCTCTTGGCTTCGGTATCAACGCTGAAACCAACCCACAGA 1669  
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540  
QY 1670 ATCTTCCTTGAGGTGATAGGACAATGTGATAGACCAGATAGAGAGCAAGCGAAGCAT 1729  
Db 541 IlePheLeuAlaGly---LysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 559  
QY 1730 TTGCAATTCCTGGGTCCGGTGAACAAGTTGAGAAGCTCATCAAAAACAGAAAGGAATCT 1789  
Db 560 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 579  
QY 1790 CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAAAGAG 1849  
Db 580 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 599  
QY 1850 TCTCCTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTCCACTCCTTTCA 1909  
Db 600 SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer 619  
QY 1910 ATTTTGAAGGCTTTTAAC 1927  
Db 620 IleLeuLysAlaPheAsn 625

Search completed: April 8, 2005, 05:01:13  
Job time : 336.773 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 8, 2005, 05:25:59 ; Search time 233.254 Seconds  
(without alignments)  
5784.401 Million cell updates/sec

Title: US-10-728-051-1  
Perfect score: 3562  
Sequence: 1 aataacatataattcattc.....cgttgtgcgtgtttctcc 2032

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 2836020

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool.p/US10728051/runat\_07042005\_125048\_14128/app.query.fasta\_1.4757  
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10728051 @CGN 1 1 663 @runat\_07042005\_125048\_14128  
-NCPUS=6 -ICPU=3 -NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=130 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pcp.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pcp.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pcp.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pcp.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pcp.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pcp.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pcp.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pcp.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pcp.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pcp.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pcp.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pcp.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	3286	92.3	626	10	US-09-847-208-28	Sequence 28, Appl
2	3286	92.3	626	14	US-10-228-806-2	Sequence 2, Appl
3	3286	92.3	626	15	US-10-100-303A-7	Sequence 7, Appl
4	3286	92.3	626	15	US-10-245-871-10	Sequence 10, Appl
5	3286	92.3	626	15	US-10-253-286-10	Sequence 10, Appl
6	3286	92.3	626	17	US-10-899-551-2	Sequence 2, Appl
7	3192	89.6	635	17	US-10-899-551-53	Sequence 53, Appl
8	3052	85.7	634	9	US-09-731-221-78	Sequence 78, Appl
9	3041	85.4	614	9	US-09-331-631A-21	Sequence 21, Appl
10	3041	85.4	614	10	US-09-847-208-27	Sequence 27, Appl
11	3041	85.4	614	14	US-10-147-095-21	Sequence 21, Appl
12	3041	85.4	614	15	US-10-100-303A-8	Sequence 8, Appl
13	1394	39.1	268	15	US-10-100-303A-55	Sequence 55, Appl
14	1282	36.0	605	15	US-10-424-599-153195	Sequence 153195,
15	1281	36.0	605	9	US-09-331-631A-25	Sequence 25, Appl
16	1281	36.0	605	14	US-10-147-095-25	Sequence 25, Appl
17	1275.5	35.8	623	15	US-10-424-599-153206	Sequence 153206,
18	1268	35.6	605	15	US-10-100-303A-110	Sequence 260105,
19	1201.5	33.7	584	15	US-10-424-599-260105	Sequence 110, App
20	1157	32.5	417	15	US-10-245-227B-14	Sequence 14, Appl
21	1157	32.5	425	15	US-10-245-227B-12	Sequence 12, Appl
22	1156	32.5	439	15	US-10-245-227B-1	Sequence 1, Appl
23	1117.5	31.4	390	15	US-10-245-227B-2	Sequence 2, Appl
24	874	24.5	324	15	US-10-425-114-44408	Sequence 4408, A
25	867	24.3	344	15	US-10-425-114-51703	Sequence 51703, A
26	865.5	24.3	666	9	US-09-331-631A-3	Sequence 3, Appl
27	865.5	24.3	666	14	US-10-147-095-3	Sequence 3, Appl
28	863.5	24.2	625	9	US-09-331-631A-5	Sequence 5, Appl
29	863.5	24.2	625	14	US-10-147-095-5	Sequence 5, Appl
30	849.5	23.8	666	9	US-09-331-631A-1	Sequence 1, Appl
31	849.5	23.8	666	14	US-10-147-095-1	Sequence 1, Appl
32	828	23.2	316	15	US-10-425-114-43905	Sequence 43905, A
33	828	23.2	316	15	US-10-425-114-49262	Sequence 49262, A
34	827	23.2	315	15	US-10-425-114-43971	Sequence 43971, A
35	816	22.9	301	15	US-10-425-114-49373	Sequence 49373, A
36	791	22.2	590	9	US-09-331-631A-8	Sequence 8, Appl
37	791	22.2	590	14	US-10-147-095-8	Sequence 8, Appl
38	790	22.2	300	15	US-10-425-114-43597	Sequence 43597, A
39	788	22.1	296	15	US-10-425-114-43984	Sequence 43984, A
40	770	21.6	291	15	US-10-425-114-51411	Sequence 51411, A
41	761	21.4	149	15	US-10-100-303A-54	Sequence 54, Appl
42	737.5	20.7	525	9	US-09-331-631A-7	Sequence 7, Appl
43	737.5	20.7	525	14	US-10-147-095-7	Sequence 7, Appl
44	732	20.6	276	15	US-10-425-114-44106	Sequence 44106, A
45	686	19.3	266	15	US-10-425-114-42946	Sequence 42946, A

ALIGNMENTS

RESULT 1  
US-09-847-208-28  
; Sequence 28, Application US/09847208  
; Publication No. US20030082190A1  
; GENERAL INFORMATION:  
; APPLICANT: Saxon, Andrew  
; APPLICANT: Zhang, Ke  
; APPLICANT: Zhu, Daocheng  
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF  
; FILE REFERENCE: IGE-MEDIATED ALLERGIC DISEASES  
; FILE REFERENCE: UC67.002A  
; CURRENT APPLICATION NUMBER: US/09/847,208  
; CURRENT FILING DATE: 2001-05-01  
; NUMBER OF SEQ ID NOS: 177  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Arachis hypogaea (Peanut)  
US-09-847-208-28  
Alignment Scores:

Pred. No.: 2,03e-261 Length: 626  
Score: 3286.00 Matches: 626  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.25% Indels: 0  
DB: 10 Gaps: 0

US-10-728-051-1 (1-2032) x US-09-847-208-28 (1-626)

QY 50 ATGAGAGGAGGTTTCTCAGTATGCTGTTCTAGGAGTCTTCTCGCTCAGTT 109  
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyLeuValLeuAlaSerVal 20

QY 110 TCTGCAACGATCCAGTATCATCCTTACAGAGAAAACAGAACCCCTGCGCCAG 169  
DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

QY 170 AGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGCATGCGAGTCT 229  
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60

QY 230 CGTGCAACGATCCAGTATGATCCTCGTTGTCTATGATCCTCGAGGACACTGCG 289  
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACCAAGTTCCCTCCAGGGGAGCGGACACAGTGGCGCCCAACCCCGGAGACTAC 349  
DB 81 ThrThrAsnGlnArgSerProProGlyGlyArgThrArgGlyArgGlnProGlyAspTyr 100

QY 350 GATGATCACCGCTCAACCCCAAGAGAGAGAGGAGCGGATGCGGACCGAGTGAACCG 409  
DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120

QY 410 AGGAGCGTGAAGAGAGAGAGTCTGAGACCAACCAAGAGAGTTCGAGGCGCACCAAGT 469  
DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140

QY 470 CATCAGACGCCAGCAAAATGAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529  
DB 141 HisGlnGlnProArgLysLysLeuArgProGluGlyArgGluGlyGluGlnTyrGlyThr 160

QY 530 CCAAGTACCATGAGGAGAGAAACATCTCGGAAACAACCTTCTACTTCCCGTCAAGG 589  
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180

QY 590 CGGTTAGCACCCCTACGGGAACCAAAACGTTAGGATCCGGTCTCGCAGAGGTTTAC 649  
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgLysValLeuGlnArgPheAsp 200

QY 650 CAAAGGTCAGGCAAGTTTCAAGATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAA 709  
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgLysValGlnLysGluAlaLys 220

QY 710 CCTAACACTTGTCTTCCCAACGACGCTGATCTGATTAACATCTCTGTTATCCAGCAA 769  
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnLysLeuValLysGln 240

QY 770 GGGCAAGCCACGTCAGCGTAGCAATGGCAATTAACAGAAAGAGCTTTTAACTCTTGACGAG 829  
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260

QY 830 GGCCATGCACTCAGAAATCCCAATCCCGTTTCAATTCCTTACATCTTGAACCGCCATGCAAC 889  
DB 261 GlyHisAlaLeuArgLysProSerGlyPheLysSerTyrLysLeuAsnArgHisAspAsn 280

QY 890 CAGAACCTCAGAGTAGCTTAAATCTCCATGCGCCGTTAAACACCCGCGCTTTCAGGAT 949  
DB 281 GlnAsnLeuArgValAlaLysLysSerMetProValAsnThrProGlyGlnPheGluAsp 300

QY 950 TTCTCCCGGCGAGCAGCGAGACCAATCATCTTACTTTCAGAGGCTTCAGCAGGAATACG 1009  
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320

QY 1010 TTGAGGCGCCCTTCAATGCGGAATTTCAATGAGATACGAGGGTGTCTGTAGAGAGAAAT 1069  
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluLysLeuArgValLeuLeuGluGluAsn 340

QY 1070 GCAGGAGTGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1129  
DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrPheThrArgSerSerGluAsn 360

QY 1130 AATGAAGAGTGTAGTCAAAAGTGTCAAAAGGAGCAGCTTGAAGAACTTACTTAAGCAGCT 1189  
DB 361 AsnGluGlyValLysValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380

QY 1190 AAATCCGTCTCAAGAAAGGCTCCGAGAGAGAGAGAGAGATATCACCACCAATCACTTG 1249  
DB 381 LysSerValSerLysLysLysGlySerGluGluGluGlyAspLysThrAsnProLysLeu 400

QY 1250 AGAGAGGCGAGCCCGATCTTCTTAAACAACTTGGAGAGTATTGAGGTCAAGCCAGCAG 1309  
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420

QY 1310 AAGAAAGAACCCCGCTTCAGGACCTGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369  
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 440

QY 1370 GGAGCTTTGATGCTCCCACTTCACTCAAGAGCCATGTTTTCGTCGTCTCAACAA 1429  
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValLysValValAsnLys 460

QY 1430 GGAATCGGAACCTTGAATCTGCTGCTGTAAGAAAGAGCAACACAGAGGAGCGGCGG 1489  
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480

QY 1490 GAAAGAGAGGAGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549  
DB 481 GluGluGluGluAspGluGluGluGluGluGluGluGluGluGluValArgArgTyr 500

QY 1550 ACAGGAGGTTCAAGAGAGGAGGATGTGTATCATGCCAGCAGCTCATCCAGTAGCCATC 1609  
DB 501 ThrAlaArgLeuLysGluGlyAspValPheLysMetProAlaAlaHisProValAlaLys 520

QY 1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTCCGTTATCAACGCTGAAACCAACACAGA 1669  
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyLysLeuAsnAlaGluAsnHisArg 540

QY 1670 ATCTTCTTTCAGGTGATAAGGACAAATGTGTATGATACACAGATAGAGAGCAAGGAT 1729  
DB 541 IlePheLeuAlaGlyAspLysAspAsnValLysAspGlnLysGlnAlaLysAsp 560

QY 1730 TTAGCATTCCTCGCTCGGTCGAGTGAAGAGTTCATCAAAACCCAGAAAGAAATCT 1789  
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuLysAsnGlnLysGluSer 580

QY 1790 CACTTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCTCGTCTCTCTGAGAAAGAG 1849  
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600

QY 1850 TCTCTGAGAAACAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909  
DB 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620

QY 1910 ATTTTGAAGGCTTTTAAAC 1927  
DB 621 IleLeuLysAlaPheAsn 626

RESULT 2  
US-10-728-806-2  
; Sequence 2, Application US/10228806  
; Publication No. US20030049237A1  
; GENERAL INFORMATION:  
; APPLICANT: Bamon, et al.  
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions  
; TITLE OF INVENTION: to Allergy  
; FILE REFERENCE: 2002834-0043

261	Db	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
890	Qy	CAGAACTCAGAGTAGCTAAATCTCATGCCCGTTAAACACACCCGCCAGTTTGAGGAT	949
281	Db	GlnAsnLeuArgValAlaIlyIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
950	Qy	TTCTTCCCGCAGCAGCCGAGACCAATCATCTACTTGCAGGGCTTTCACAGGAATACG	1009
301	Db	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
1010	Qy	TTTCGAGGCCCTTCAATCGGGAATTCAATGACATACGGAGGCTGCTGTGTAGAGAGAAT	1069
321	Db	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuGluGluAsn	340
1070	Qy	GCAGGAGGTGAGCAACAGGAGACAGGCGCAGAGCGGATGAGTAGTACTCGAGCTAGTGAGAAC	1129
341	Db	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360
1130	Qy	AATGAAGGAGTAGTAGTCAAAAGTGTCAAAGGACAGCGTTGAAGAACTTACTTAAGACACGCT	1189
361	Db	AsnGluGlyValIleValIyValSerLysGluHisValGluGluLeuThrLysHisAla	380
1190	Qy	AAATCCGTCTCAAGAAAGGCTCCGAAGAAGAGAGGGAGATATCAACCAACCCCAATCAACTTG	1249
381	Db	LysSerValSerLysIySerGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
1250	Qy	AGAGAAGGCAGGCCGATCTTTCTAAACAACCTTTGGGAAGTTATTGTAGGTGTAAGCCAGAC	1309
401	Db	ArgGluGlyGluProAspLeuSerAsnAsnPhgIyLysLeuPheGluValLysProAsp	420
1310	Qy	AAGAAGAACCCCCAGCTTTCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAAAGAA	1369
421	Db	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
1370	Qy	GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCCATGGTTATTCGTGCTGCTCAACAAA	1429
441	Db	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValLysLys	460
1430	Qy	GGAACTTGGAAACCTTGAACTCTGGTGGCTGTAAAGAAAAGAGCAACAACAGAGGGGACGGCGG	1489
461	Db	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	480
1490	Qy	GAAGAAGAGGAGCAGAAAGCAGGAAGAGGAGGGAAGTAAACAGAGGTGCGTAGGTAC	1549
481	Db	GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
1550	Qy	ACAGCAGGTTGAAGGAGCGATGTGTTCAATCATGCCACAGCTCATCCAGTAGCCATC	1609
501	Db	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
1610	Qy	AACGCTTCCTCCGAACTCCATCTGCTGGCTTCGGTATCAACGCTGAAAAACAACACAGA	1669
521	Db	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
1670	Qy	ATCTTCTCTTCGAGGTGATTAAGGACAATGTGATAGACCAGATAGAGAAGCAGCAAGGAT	1729
541	Db	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
1730	Qy	TTTAGCATTCCTCGGTCGGGTGAACAAGTTTGAGAAGCTCATCAAAAACCAAGAGGAATCT	1789
561	Db	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
1790	Qy	CACCTTGTGAGTGCTCGTCTCTCAATCTCAATCTCCGTGCTCTCCTGAGAAAAGAG	1849
581	Db	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
1850	Qy	TCTCTCTGAGAAAGAGGATCAAGAGGAGGAAAAACAAGAGGGAAGGCTCCACTCTTTCA	1909
601	Db	SerProGluLysGluAspGlnGluGluGluGluGluGluGluGlyGlyGlyProLeuLeuSer	620
1910	Qy	ATTTTGAAGGCTTTTAACT	1927

```
Db      621 IleLeuLysAlaPheAsn 626
RESULT 3
US-10-100-303A-7
; Sequence 7, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo-Ara h 1
US-10-100-303A-7
Alignment Scores:
Pred. No.: 2,03e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 15 Gaps: 0
US-10-728-051-1 (1-2032) x US-10-100-303A-7 (1-626)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGATCCTTGTCCTGGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCCAAGTATCATCACTTACAGAGAAACAGAGAACCCCTCGGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTCTCAACAGGACCGGATGCTTGAAGCAAAAGGATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGlnProAspAspLeuLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTTCGAGTATGATTCCTCGTTGTCTATGATCCTCGAGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACACCAACCAAGTTCCTCCCTCAGGGAGGGACACGTGGCCGCCCAACCGGAGCTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCTCAACCCGAGAGAGAGAGGAGGCGCATGGGACACGATGGACCG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrProAlaGlyPro 120
QY 410 AGGGAGCGTGAAGAGAGAGAGTGGAGACCAACCAAGAGAGATTCGAGCGGACCAAGT 469
Db 121 ArgGluArgGluArgGluLysTyrArgGlnProArgGluAspTyrArgGlyProSer 140
QY 470 CATCAGACGCCAGGAATAGCCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnTyrGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGAGAGAGACATCTCGGAACCAACCTTTCTACTTCCCTCAAGG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACCGTAGGATCCGGGTCCTGCAGAGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGCGAGTTTCAGAACTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAA 709
Db 110 CCTAACACTCTTGTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATTATCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTACCGTAGCAAAATGCAATTAACAGAAAGAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACTCAGAAATCCCATCCGTTTCATTTCTTACATCTTGAACCGCATCAGAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCGTTTACACACACCCGGCCAGTTTGAGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCAGGAGCAGCCGAGACCAATCATCTTCTTCTGAGGGCTTCAGAGGAATACG 1009
Db 301 PhePheProAlaSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGGCGCTTCAATGCCGAATTCATGAGATACGAGGGTCTGTTAGAGAGAGAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGGAGGCGAGCTGAGTACTCGAGTAGTAGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
QY 1130 AATGAGAGGTATAGTCAAGTGTCAAGAGGACGTTGAAGAACTTACTTAAGCACGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCGCTCTCAAGAAAGGCTCCGAGAGAGGAGAGATATCACCAACCCCAATCACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAGGCGAGCCCGATCTTTTCAACAACTTTGGGAGTATTATTTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAGAACCCCGACCTTCAGGACCTGGACATGATGCTCACCTGTGTGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluLysGlu 440
QY 1370 GGAGCTTTTCATGCTCCCACTTCAACTCAAGGCGATGTTTATCTCGTCTCAACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAATCGGAACTTGAATCTCGGTGTAAAGAAAGAGCAACACAGAGGGCGGCGG 1489
Db 461 GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGGAGCAACAGCAAGAGAGAGGAGGAGTAAACAGAGAGTGGCTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCAGGTGTGAAGGAGCGGATGTGTTCATCATGCCAGCGCTCATCCAGTGCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AAGCTTCTCCGAACTCCATCTGCTTGGTTCGTTATCAACGCTCAAAACACCAACAG 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTGCAGGTGATAAGGACAAATGTGATAGACCATAGAGAGAGCAAGCGAGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTTCCCTGGGTGAGCAAGTTGAGAGAGCTCATCAAAACACAGAGGATCTT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
```

QY 1790 CACTTGTGTGCTCTCTCAATCTCAATCTCAATCTCCGTCTCTCTGAGAAAGAG 1849  
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600  
QY 1850 TCTCTCAGAAAGAGGATCAAGAGGAGAAACCAAGAGGAGGAGGTCCTCACTCTTCA 1909  
DB 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyLysGlyProLeuLeuSer 620  
QY 1910 ATTTGAAGGCTTTTAAAC 1927  
DB 621 IleLeuLysAlaPheAsn 626

RESULT 4  
US-10-245-871-10  
; Sequence 10, Application US/10245871  
; Publication No. US20030235594A1  
; GENERAL INFORMATION:  
; APPLICANT: HUMPHREYS, ROBERT  
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES  
; FILE REFERENCE: REH-2013  
; CURRENT APPLICATION NUMBER: US/10/245,871  
; CURRENT FILING DATE: 2003-01-09  
; PRIOR APPLICATION NUMBER: 10/197,000  
; PRIOR FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: 09/396,813  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 905  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Arachis hypogaea  
US-10-245-871-10

Alignment Scores:  
Pred. No.: 2,03e-261 Length: 626  
Score: 3286.00 Matches: 626  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.25% Indels: 0  
DB: 15 Gaps: 0

US-10-728-051-1 (1-2032) x US-10-245-871-10 (1-626)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTCTGAGGATCTCTGCTGGCTTCAGTT 109  
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20  
QY 110 TCTGCAACGATGCCAAGTCAATCACCTTACCAGAAGAAACAGAGAACCCCTCGCCCCAG 169  
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40  
QY 170 AGTGCTCCAGATGTTCTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGAGTCT 229  
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60  
QY 230 CGCTGCACCAAGCTCCGAGTATGATCCTCGTTGTCTCATGATCTCGAGGACACTGGC 289  
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80  
QY 290 ACCACCAACCAAGCTTCCCTCCAGGGGAGCGGACACGTCGGCCGCCAACCCGGAGACTAC 349  
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100  
QY 350 GATGATACCGCGCTCAACCCGGAAGAGAGAGAGGAGCGGATGGGACCGAGTGGACCG 409  
DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120  
QY 410 AGGAGCGCTCAAGAGAGAGACTCGAGACACCAAGAGAGAGATTTGGAGCCCAAGT 469  
DB 121 ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140

QY 470 CATCAGCAGCCACGGAATAAGCCCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529  
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr 160  
QY 530 CAGGTAGCCATGTGAGGAG 589  
DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180  
QY 590 CGGTTTAGCACCCGCTACGGGAACCAACCGTAGGATCCGGTCCCTCGAGAGGTTTGAC 649  
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgLleArgValLeuGlnArgPheAsp 200  
QY 650 CAAAGGTCAAGGCAGTTTCCAGAAATCCAGAAATCACCGTATTGTCGAGATCGAGGCCAA 709  
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220  
QY 710 CTTAAACACTCTTGTCTTCCCAAGCAGCTGATGCTGATAAACATCTTGTATTCAGCAA 769  
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240  
QY 770 GGGCAAGCCACCGTAGCCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG 829  
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260  
QY 830 GGCCATGCACTCAGAAATCCCATCCCGTTCATTCTCTACATCTTGAACCGCCATGCAAC 889  
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280  
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCCGCCAGTTTGAGGAT 949  
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300  
QY 950 TTCTTCCCGCGAGCAGCAGACCAATCATCTCTCTGAGGAGCTTCAGCAGAGATACG 1009  
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320  
QY 1010 TTGGAGCGCGCTTCAATCGGAAATTCATAGATACGAGGAGGCTGTGTAGAGAGAGAT 1059  
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340  
QY 1070 GCAGGAGGTGAGCAAGCAGAGCAGGCGCAGGCGATGAGTACTCGAGTAGTAGAAC 1129  
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgTyrSerThrArgSerSerGluAsn 360  
QY 1130 AATGAAGGAGTGTAGTCAAAAGTGTCAAAGGACGACGTTGAAGAACTTACTTAAGACGCT 1189  
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380  
QY 1190 AATCCGTCTCAAAGAAAGGCTCCGAAAGAGGAGGAGATATCAACCAACCAACTTG 1249  
DB 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400  
QY 1250 AGAAGAGGCGCCGATCTTTCTAACTTTGGGAAGTATTATTCAGGTGAAGCCAGAC 1309  
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420  
QY 1310 AGAAGAAACCCCGAGCTTCAGGACCTCGGACATGATCTCACCTGTGTAGAGATCAAGAA 1369  
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440  
QY 1370 GGAGCTTTGATGTCCTCCACACTTCAACTCAAAGGCCATGTTATCGTCTGCTCAACAAA 1429  
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460  
QY 1430 GGAACCTGGAAACCTTGAATCTGTGGCTGTAAAGAAAGAGCAACACAGAGGAGCGCGG 1489  
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480  
QY 1490 GAAGAGAGGAGGAGCAACAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1549  
DB 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500





|||||  
421 LysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440  
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATCGTCGTCGTCACAA 1429  
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460  
QY 1430 GGAACCTGGAACCTTGAACCTCGCTGCTGTGAAGAAAGAGCAACACAGAGGGACGGCG 1489  
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480  
QY 1490 GAAGAAGAGGAGGACGAGACGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1549  
Db 481 GluGluGluGluAspGluAspGluGluGluGluGluGluGluGluGluGluGluGluGlu 500  
QY 1550 ACAGCGAGGTGAAGGAGCGATGCTTCATCATGCGCAGCAGCTCATCCAGTAGCCCATC 1609  
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520  
QY 1610 AAGCTTCCCGCACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAGA 1669  
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540  
QY 1670 ATCTTCTTCAGTGATAGGACAATGTGATAGCAGATAGAGAGCAAGCGAAGAT 1729  
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560  
QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAGTTGAGAAAGCTCATCAAAACCAAGAAGGAATCT 1789  
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580  
QY 1790 CACTTTGTAGTGCTGCTCTCAATCTCAATCTCAATCTCGCTGCTCTCTGAGAAAGAG 1849  
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600  
QY 1850 TCTCCTCAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1909  
Db 601 SerProGluLysGluAspGlnGluGluGluGluGluGluGluGluGluGluGluGluGlu 620  
QY 1910 ATTTTGAAGGCTTTTAAC 1927  
Db 621 IleLeuLysAlaPheAsn 626

RESULT 6  
US-10-899-551-2  
; Sequence 2, Application US/10899551  
; Publication No. US20050063994A1  
; GENERAL INFORMATION:  
; APPLICANT: Caplan, Michael J.  
; APPLICANT: Burks, A. Wesley  
; APPLICANT: Sampson, Hugh A.  
; APPLICANT: Howard, Sosin B.  
; APPLICANT: Bottomly, Kim H.  
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy  
; FILE REFERENCE: 2002834-0233  
; CURRENT APPLICATION NUMBER: US/10/899,551  
; CURRENT FILING DATE: 2004-07-26  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: species Arachis hypogaea  
US-10-899-551-2

Alignment Scores:  
Pred. No.: 2,03e-261 Length: 626  
Score: 3286.00 Matches: 626  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.25% Indels: 0  
DB: 17 Gaps: 0

US-10-728-051-1 (1-2032) x US-10-899-551-2 (1-626)  
QY 50 ATCAGAGGAGGGGTTTTCTCCACTGATGCTGTTCTAGGGATCCTTGTCTCGGTTCAGTT 109  
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20  
QY 110 TCTGCAACCGCATCCCAAGTCATCACCTTACCAGAAGAAACAGAGAACCCCTCGGCCAG 169  
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysThrGluAsnProCysAlaGln 40  
QY 170 AGTGCTCTCCAGAGTTGTCAACAGGAACCGGATGACTTTGAAGCAAAAGGATCGAGTCT 229  
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspLeuLysGlnLysAlaCysGluSer 60  
QY 230 CGTGCAACCAAGCTCAGTATGATCTCTGTTGTCTATGATCCTCGAGACACACTGGC 289  
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80  
QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGACACGTCGGCCGCCAACCCGGAGACTAC 349  
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100  
QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGGAGGCGGATGGGGACCAAGTGGACCG 409  
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120  
QY 410 AGGAGCGTGAAGAAGAAGACTCGAGACAAACCAAGAGAAGATTGGAGCGACCAAGT 469  
Db 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140  
QY 470 CATCAGCAGCCGGAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529  
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTyrGlyThr 160  
QY 530 CCAAGTAGCATCTGAGGAGAGAAACATCTCGAAACAAACCTTTCTACTTCCGTCAAGG 589  
Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheThrProSerArg 180  
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACCGTAGGATCCGGTCTCTGACAGAGTTTAC 649  
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValGlnIleGluAlaLys 200  
QY 650 CAAAGGTCAAGCGAGTTTCAGAAATCCAGAAATCAGGTTATGTCAGATCGAGGCCAAA 709  
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220  
QY 710 CCTAACACTCTGCTTCTCCCAAGCAGCTGATGCTGATAACATCCTTGTATTATCCAGCA 769  
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240  
QY 770 GGGCAAGCCACCGTAGCCGTCAGCAATGGCAATACACAGAAAGAGCTTTAATCTTGACG 829  
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260  
QY 830 GGCATGCACTCAGATCCCATCCGTTTCATTTCTACATCTTGAACCCGCATGACAAC 889  
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280  
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCGGCCAGTTTGAAGAT 949  
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300  
QY 950 TTCTTCCCGCGAGCAGCGAGACCAATCACTCTTCTGAGGGCTTTCAGCAGGAATACG 1009  
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320  
QY 1010 TTGAGGCGCCCTTCAATGCGGAATTCATGAGATACGAGGCTGCTGTTAGAGAGAAAT 1069  
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340  
QY 1070 GCAGAGGTGACGACAGAGGAGAGGCGAGCGATGGAGTACTCGAGTAGTGAGAAC 1129  
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGluArgGlyArgGlyArgGlyArgGlyArgGly 360



```
Db 297 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 316
Qy 1010 TTGGAGCGCCCTTCAATCGGAATTCATAGATACGGAGGCTGCTGTAGAGAGAAT 1069
Db 317 LeuGluAlaAlaPheAsnAlaGluPheAsnGluLeuArgValLeuLeuGluAsn 336
Qy 1070 GCAGGAGTGCAGCAGAGGAGAGAGCGGAGCGGATGGAGTACTCGAGTAGTAGAAC 1129
Db 337 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 356
Qy 1130 AATGAAGGAGTGTAGTCAAGGTGTCAAAGGACGCTTCAAGCACTTACTTAAGCAGCT 1189
Db 357 AsnGluGlyValLeuValValSerLysGlnHisValGluLeuThrLysHisAla 376
Qy 1190 AAATCCGTCTCAAGAAAGCTCCGAAAGAGGAGATATACCAACCCCAATCAACTTG 1249
Db 377 LysSerValSerLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 396
Qy 1250 AGAAGGCGAGCCGATCTTTCTAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 397 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 416
Qy 1310 AAGAAGACCCCGCTTCAGGACCTCGAGCATGCTCACCTGCTAGAGATCAAGAA 1369
Db 417 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 436
Qy 1370 GGAGCTTTGATGCTCCACACTTCACTCAAGGCCATGTTATCGTCGTCAACAAA 1429
Db 437 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 456
Qy 1430 GGAATCGAAACCTTGTAACCTCGTGGCTTGAAGAAAGAGCAACAAGAGGGGCGCGG 1489
Db 457 GlyThrGlyAsnLeuGluLeuValAlaValLargLysGluGlnGlnArgGlyArg 476
Qy 1490 GAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1549
Db 477 GluGluGluGluAspGluAspGluGluGluGluGluGluGluGluGluGluGlu 496
Qy 1550 ACAGCAGGTTGAAGGAGCGATGCTTTCATCATGCGCAGCATCATCCAGTAGCCATC 1609
Db 497 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAla 516
Qy 1610 AACGTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCCAGA 1669
Db 517 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyLeuAsnAlaGluAsnAsnHisArg 536
Qy 1670 ATCTTCTTCAGGTGATAGGACATGTGATAGACCATAGAGCAGGAGCAGGAGGAGGAT 1729
Db 537 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 556
Qy 1730 TTAGCATTCCTGGGTCCGCTGAACAAAGTTGAGAAGCTCATCAAAACCAAGAGGAATCT 1789
Db 557 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 576
Qy 1790 CACTTGTGAGTCTCTCTCAATCTCAATCTCAATCTCCGCTGCTCTCTGAGAAAGAG 1849
Db 577 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 596
Qy 1850 TCTCCTGAGAGGAGGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1909
Db 597 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyLysGlyProLeuLeuSer 616
Qy 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 617 IleLeuLysAlaPheAsn 622
```

## RESULT 8

```
US-09-731-221-78
; Sequence 78, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
```

```
; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
; OTHER INFORMATION: Hypogaea
US-09-731-221-78
```

## Alignment Scores:

Pred. No.:	3,79e-242	Length:	634
Score:	3052.00	Matches:	584
Percent Similarity:	96.86%	Conservative:	3
Best Local Similarity:	96.37%	Mismatches:	19
Query Match:	85.68%	Indels:	0
DB:	9	Gaps:	0

US-10-728-051-1 (1-2032) x US-09-731-221-78 (1-634)

```
Qy 110 TTGCAACGCATGCCAAGTCATCACCTTACCGAGAGAAACAGAGAACCCCTCGGCCAG 169
Db 16 SerSerThrHisAlaLysSerProTyrGlnAlaLysThrGluAsnProCysAlaGln 35
Qy 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGCATCGAGTCT 229
Db 36 ArgCysLeuGlnSerCysGlnGluProAspAlaLeuLysGlnLysAlaCysGluSer 55
Qy 230 CGTGCACCAAGTCCAGGTATGATCCTCGTTGTCTATCATCTCGAGGACACACTGGC 289
Db 56 ArgCysThrLysLeuGluTyrAspProArgCysAlaTyrAspProArgGlyHisThrGly 75
Qy 290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACACAGTGGCCGCCCAACCCGGAGACTAC 349
Db 76 ThrThrAsnGlnArgSerProGlyGluAlaThrArgGlyArgGlnProGlyAspTyr 95
Qy 350 GATGATGACCGCGTCAACCCCGAGAGAGGAGGCGCGATGGGGACCAGCTGGACCG 409
Db 96 AspAspAlaArgArgGlnProArgAlaGluGluGlyGlyArgTrpGlyProAlaGlyPro 115
Qy 410 AGGAGCGTGAAAGAGAGAGAGACTCGAGACAACCAAGAGAGATTTCGAGGCGCACCAAGT 469
Db 116 ArgGluArgGluArgGluGluAspAlaArgGlnProArgGluAspTrpAlaArgProSer 135
Qy 470 CATCAGCAGCCACGGAAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 136 HisGlnGlnProArgLysAlaArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 155
Qy 530 CAGGTAGCAGTCAGGAGGAGAGAACATCTCGGAACAACCTTCTTACTTCCCGTCAAGG 589
Db 156 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 175
Qy 590 CGGTTTAGCACCCGCTACCGGAACCAAAACGGTAGGATCCGGTCTCTCGAGAGGTTTGC 649
Db 176 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 195
Qy 650 CAAAGTCAAGGAGTTTCAGAAATCTCCAGAAATCACCGTATTCGTGAGATCGAGGCCAAA 709
Db 196 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 215
Qy 710 CCTAACACTCTTGTCTTCTCCAGAGCAGCTGATGCTGATAAACATCTTGTATTATCCAGCAA 769
Db 216 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 235
Qy 770 GGGCAAGCCACCGTGAACCGCTAGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGG 829
Db 236 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 255
```



QY 590 CGGTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCGCAGAGGTTGAC 649  
 Db 175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgValLeuGlnArgPheAsp 194  
 QY 650 CAAAGCTCAAGGAGTTCAGAACTCCAGAAACACCGTATTGTCGAGATCCAGGCCAAA 709  
 Db 195 GlnArgSerGlnPheGlnAsnLeuGlnAsnHisArgValGlnGlnLeuAlaArg 214  
 QY 710 CCTAAACACTCTGCTCTTCCCAAGCAGCTGATGCTGATAACATCTGTTATCCAGCAA 769  
 Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAsnLeuValLeuGlnGln 234  
 QY 770 GGCAAGCCACCTGACCGTAGCAAAATGCAATAACAGAAAGCTTTAATCTTGACGAG 829  
 Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuLeuAspGlu 254  
 QY 830 GGCATGCTCACTCAGATCCCATCCGTTTCAATCTTCTACATCTTGACCGCATGACAC 889  
 Db 255 GlyHisAlaLeuArgLysProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274  
 QY 890 CAGAACCTCAGAGTAGTAAATCTCCATCCCGCTTAAACACACCCGCCAGTTTGAGGAT 949  
 Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294  
 QY 950 TTCTTCCCGCGAGCAGCCAGACCAATCATCTCTTCTGAGGCTTCAGCAGGAATACG 1009  
 Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314  
 QY 1010 TTGAGCCGCTTCAATCGGAATTCATAGATACGAGGCTGCTGTTAGAAAGAAAT 1069  
 Db 315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 334  
 QY 1070 GCAGGAGGTGACGAAGAGGAGAGGCGAGCGATGGAGTACTCGGAGTAGTGAAC 1129  
 Db 335 AlaGlyGlyGlnGlnGluArgGlyGlnArgSerThrArgSerSerAsp--- 353  
 QY 1130 AATCAAGAGGTGATGATCAAGTGTCAAAGGACGCTTCAAGAACTTACTAAGACCGCT 1189  
 Db 354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373  
 QY 1190 AATCGCTCAAGAAAGCTCGAAGAGGAGGAGATATACCAACCCCAATCAACTTG 1249  
 Db 374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392  
 QY 1250 AGAAGAGCCAGCCGATCTTCTCAACACTTTGGGAAGTATTGAGGTGAAGCCAGAC 1309  
 Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412  
 QY 1310 AAGAAGAACCCCGCTTCAGGACCTGGACATGCTCACTGTGTAGAGATCAAGAA 1369  
 Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432  
 QY 1370 GAGCTTTGATGCTCCACACTTCAACTCAAGCCGATGTTATCGTGTGCTCAACAA 1429  
 Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452  
 QY 1430 GGAACCTGGAACCTGGAACCTGCTGGCTGAAGAAAGAGCAACACAGAGGGGACGCGG 1489  
 Db 453 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGlnGlnArgGlyArgArg 472  
 QY 1490 GAA-----GAAGAGGAGGACGAACGAAGAGAGGAGGAGGAGTAAACAGAGAGGTG 1540  
 Db 473 GluGlnGlnLysTrpGluGluGluGluAspGluGluGluGlySerAsnArgGluVal 492  
 QY 1541 CGTAGGTACACAGCGAGTTGAAGAGGCGATGTTTCATCATGCGAGCGAGTCTATCCA 1600  
 Db 493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512  
 QY 1601 GTAGCCATCAACGCTTCTCGAATCTCATCTGCTGGCTTCGTTATCAACGCTGAAC 1660  
 Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532

QY 1661 AACCAAGAAATCTTCTTCAGGTGATAGGACAATGTGATAGACGAGATAGAGAACAA 1720  
 Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552  
 QY 1721 GCGAAGGATTAGCATCTCCCTCGGTGGAACAAGTTGAGAAGCTCATCAAAACCAG 1780  
 Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572  
 QY 1781 AAGGAATCTCACTTTGAGTGTCTCTCTCAATCTCAATCTCAATCTCTCTCTCTCT 1840  
 Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588  
 QY 1841 GAGAAAGATCTCTGAGAAAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1900  
 Db 589 -----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyPro 605  
 QY 1901 CTCCTTTCAATTTGAAGGCTTTTAAAC 1927  
 Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614

RESULT 10  
 US-09-847-208-27  
 ; Sequence 27, Application US/09847208  
 ; Publication No. US20030082190A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saxon, Andrew  
 ; APPLICANT: Zhang, Ke  
 ; APPLICANT: Zhu, Daocheng  
 ; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF  
 ; TITLE OF INVENTION: IGH-MEDIATED ALLERGIC DISEASES  
 ; FILE REFERENCE: UC67.002A  
 ; CURRENT APPLICATION NUMBER: US/09/847,208  
 ; CURRENT FILING DATE: 2001-05-01  
 ; NUMBER OF SEQ ID NOS: 177  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 27  
 ; LENGTH: 614  
 ; TYPE: PRT  
 ; ORGANISM: Arachis hypogaea (Peanut)  
 US-09-847-208-27

Alignment Scores:  
 Pred. No.: 3,01e-241 Length: 614  
 Score: 3041.00 Matches: 595  
 Percent Similarity: 96.03% Conservative: 9  
 Best Local Similarity: 94.59% Mismatches: 7  
 Query Match: 85.37% Indels: 18  
 DB: 10 Gaps: 7

US-10-728-051-1 (1-2032) x US-09-847-208-27 (1-614)

QY 50 ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTGCTAGGATCTCTGCTCGCTTCAGTT 109  
 Db 1 MetArgLysArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20  
 QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACAGAGAAACACAGAACCCCTCGCCAG 169  
 Db 21 SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln 38  
 QY 170 AGGTGCTCCAGATGTTGTCAACAGGAACCGGATGACTTGAGCAAAAGGATCGAGTCT 229  
 Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58  
 QY 230 CGGTGCACCAAGCTCCAGTATGATCTCGTTGTGTATGATCTCTCGAGGACACACTGGC 289  
 Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74  
 QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGCGGACACGTCGGCCGCCAACCCCGAGACTAC 349  
 Db 75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94  
 QY 350 GATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGGATGGGACCGAGTGGACCG 409

Db 95 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGluPro 114  
QY 410 AGGAGCGTCAAGAGAGAAAGAACTCGAGACAACCAAGAGAAATTTGGAGGCGACCAAGT 469  
Db 115 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer 134  
QY 470 CATCAGAGCCAGCGAAATAAGGCCCGAAGAGAGAGAGAGAAACAAGAGTGGGAAACA 529  
Db 135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTrpGlyThr 154  
QY 530 CCAGGTAGCATGTAGGAGAGAAACATCTCGAACAACCCCTTCTACTTCCGTCGAAG 589  
Db 155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheThrPheProSerArg 174  
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACCGTAGGATCCGGTCCCTCGACAGGTTTGAC 649  
Db 175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194  
QY 650 CAAAGGTCAAGCGATTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAA 709  
Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214  
QY 710 CTTAAACACTCTGTTCTCCAGACGCTGATGCTGATAACATCTTGTATCCAGCAA 769  
Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234  
QY 770 GGGCAAGCCACCGCTAGCAATGCAATTAACAGAAAGACGTTTAACTCTTGACGAG 829  
Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254  
QY 830 GGCATGCTCAATCCCATCCGTTTCAATTCCTACATCTTGAAACCGCCCATGACAAC 889  
Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274  
QY 890 CAGAACTCAGAGTACGTAATCTCCATGCCGCTTAAACACCCCGCCAGTTTGGAGT 949  
Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294  
QY 950 TTCTTCCCGCAGCAGCCGAGACCAATCATCTCTGAGGCGTTCAGCAGGAATACG 1009  
Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314  
QY 1010 TTGGAGCGCCTTCAATCGGAATTCATAGATACGAGAGGTGCTGTTAGAGAGAAAT 1069  
Db 315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 334  
QY 1070 GCAGGAGGTGAGCAGAGGAGAGCGCAGAGCGATGAGTACTCGAGTACTGAGAAC 1129  
Db 335 AlaGlyGlyGlnGluGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353  
QY 1130 AATCAAGGAGGTAGTCAAGTGTCAAAGGACGCTTGAAGAACTTACTAAGCAGCT 1189  
Db 354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373  
QY 1190 AATCCGTCTCAAAGAAAGCTCCGAGAGAGGAGATATCAACCAACCCCAATCAACTTG 1249  
Db 374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392  
QY 1250 AGAAGAGCGAGCCGATCTTTCTAACTTTGGGAAGTATTTCAGGTGAGCCAGAC 1309  
Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412  
QY 1310 AAGAAGAACCCCGAGTTCAAGACCTGGACATGATGCTACCTGTGTAGAGATCAAGAA 1369  
Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432  
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTGTGTCGTCACAA 1429  
Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 452  
QY 1430 GGAACGTGAACCTGCACTCGTGGCTGAAGAAAGAGCAACAACAGAGGCGGCGCG 1489  
Db 453 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGlnGlnGlnArgGlyArgArg 472

QY 1490 GAA-----GAAGAGGAGGACGACGAGAGAGAGGAGGAGTAAACAGAGAGGTG 1540  
Db 473 GluGlnGluTrpGluGluGluGluAspGluGluGluGlySerAsnArgGluVal 492  
QY 1541 CGTAGGTACACAGCGGTTGAAGAGCGGATGTTTCATCATGCCAGCAGCTCATCCA 1600  
Db 493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512  
QY 1601 GTAGCCATCAACGCTTCTCCGAATCCATCTGCTTGGTTCGGTATCAACGCTGAAAC 1660  
Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532  
QY 1661 AACCCAGAGATCTTCTTTCAGGTGATAAGCAATGTGATAGACAGATAGAGAGCAA 1720  
Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552  
QY 1721 GCGAGGATTTAGCATTCCTCGGTCCGGTGAACAGTGAAGCTCATCAAAACCAAG 1780  
Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572  
QY 1781 AAGGAATCTCACTTGTGAGTCTGCTCTCAATCTCAATCTCAATCTCCGTCGTCCT 1840  
Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588  
QY 1841 GAGAAAGATCTCTCAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGTCCA 1900  
Db 589 -----SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyPro 605  
QY 1901 CTCCTTTCATTTTGAAGGCTTTTAAAC 1927  
Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614

## RESULT 11

US-10-147-095-21  
; Sequence 21, Application US/10147095  
; Publication No. US2003017127A1  
; GENERAL INFORMATION:  
; APPLICANT: Manners, John M.  
; APPLICANT: Marcus, John Paul  
; APPLICANT: Goulter, Kenneth C.  
; APPLICANT: Green, Jodie L.  
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS  
; FILE REFERENCE: CULLN23.001APC  
; CURRENT APPLICATION NUMBER: US/10/147,095  
; CURRENT FILING DATE: 2002-05-15  
; PRIOR APPLICATION NUMBER: US/09/331,631A  
; PRIOR FILING DATE: 1999-06-21  
; PRIOR APPLICATION NUMBER: PCT/AU97/00874  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: AU PO 4275  
; PRIOR FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 21  
; LENGTH: 614  
; TYPE: PRT  
; ORGANISM: Peanut  
US-10-147-095-21

Alignment Scores:  
Pred. No.: 3,01e-241 Length: 614  
Score: 3041.00 Matches: 595  
Percent Similarity: 96.03% Conservative: 9  
Best Local Similarity: 94.59% Mismatches: 7  
Query Match: 85.37% Indels: 18  
DB: 14 Gaps: 7

US-10-728-051-1 (1-2032) x US-10-147-095-21 (1-614)

QY 50 ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTGTAGGGATCCTTGTCTCGCTTCAGTT 109  
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20





Score:	3041.00	Matches:	595
Percent Similarity:	96.03%	Conservative:	9
Best Local Similarity:	94.59%	Mismatches:	7
Query Match:	85.37%	Indels:	18
DB:	15	Gaps:	7

US-10-728-051-1 (1-2032) x US-10-100-303A-8 (1-614)

QY	50	ATGAGAGGAGGGTTTCTCCACTGATGCTGTCTAGGATCCTTGTCTGGCTTCAGTT	109
DB	1	MetArgLysValSerProLeuMetLeuLeuGlyIleuValLeuAlaSerVal	20
QY	110	TTGCAACCGCATCCAAAGTTCATCCTTACCAAGAAACAGAGAACCCCTCGCCACG	169
DB	21	SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln	38
QY	170	AGTGCTCCAGATGTCTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGAGTCT	229
DB	39	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	58
QY	230	CGCTGCACCAAGTCCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACCTGGC	289
DB	59	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly	74
QY	290	ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTCGGCCGCCAACCCCGAGACTAC	349
DB	75	AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	94
QY	350	GATGATGACCGCGCTCAACCCCAAGAGAGAGGAGGCGGATGGGGACCGACTGACCG	409
DB	95	AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGluPro	114
QY	410	AGGAGCGTCAAGAGAGAAAGACTCGAGACAACCAAGAGAAAGATTGGAGGCGCAACAGT	469
DB	115	ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer	134
QY	470	CATCAGCAGCACCGAAATAAGCCCGCAAGAGAGAAAGAGAAACCAAGATGGGAAACA	529
DB	135	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr	154
QY	530	CCAGGTAGCATGTGAGGAGAAACATCTCGAAACAACCCCTTCTACTTCCCGTCAAG	589
DB	155	ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	174
QY	590	CGGTTTAGCACCCGCTACGGGAACCAACCGTAGGATCCGGTCTCGCAGAGGTTTGAC	649
DB	175	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	194
QY	650	CAAGGTCAAGCGAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGAGATCGAGCCAAA	709
DB	195	GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg	214
QY	710	CCTAACACTTGTCTTCTCCAGCAGCGTGTGATGATCAACATCCTTGTATCCAGCAA	769
DB	215	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	234
QY	770	GGCAAGCCACCGTGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGAACGAG	829
DB	235	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	254
QY	830	GGCCATGCATCAGATCCCATCCCGTTTCATTTCTTACATCTTGAACCCGCCATGACAA	889
DB	255	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	274
QY	890	CAGAACTCAGATAGTAAATCTCCATGCCCGCTTAAACACACCCCGCCAGTTTGAAGAT	949
DB	275	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	294
QY	950	TTCTTCCCGCGCAGCGCAGACCAATCATCTTCTGAGGCGCTTCAGCAGGAATACG	1009
DB	295	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	314
QY	1010	TTGGAGCGCGCTTCAATCGGAAATTCATAGATACGGAGGTGCTGTTAGAAGAGAAT	1069

RESULT 13

US-10-100-303A-55  
; Sequence 55, Application US/10100303A  
; Publication No. US20030202980A1  
; GENERAL INFORMATION:  
; APPLICANT: Caplan, et al.  
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction  
; TITLE OF INVENTION: to Allergy  
; FILE REFERENCE: 2002834-0166  
; CURRENT APPLICATION NUMBER: US/10/100,303A

Qy	50	ATGAGAGGGAGGGTTTCTCCACTGATGCTGTGTAGGGATCCCTTGTCTCGCTTCAGTT	109
Db	2	MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal	19
Qy	110	TCGTCAACGCGATGCCAAGTCATCACCTTACCAGAGAAGAAACAGAGAACCCCTGCGCCACG	169
Db	20	SerValSerPheGlyIleAla-----TyrTrpGluLys-----GluAsnProLysHisAsn	36
Qy	170	AGGTGCTCTCCAGAGTTGTCAAACGAGAACCGGATGACTTGAAGAAAAAGCATGCGAGTCT	229
Db	37	LysCysLeuGlnSerCysAsnSerSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla	56
Qy	230	CGTGTCAACCAAGCTCGAGTATGAT-----CCTCGTTGT	262
Db	57	ArgCysAsnLeuLeuLysValGluLysGluCysGluGluGlyGluIleProArg---	75
Qy	263	GTCATGATCTCGAGGACACACTGGGCACCAACCAACCGTTCCTCCAGGGAGCGG	322
Db	76	-----ProArgProArg-----ProGlnHisProGlu	84
Qy	323	ACAGTGGCGCCCAACCGGAGACTACGATGATGAC-----	358
Db	85	ArgGluProGlnProGlyGlyLysGluGluAspGluAspGluGlnProArgProIle	104
Qy	359	-----CGCGCTCAACCCGGAAGAG-----GAAGGAGGCCGA	391
Db	105	ProPheProArgProGlnProArgGlnGlnGluHisGluGlnArgGluGlnGlu	124
Qy	392	TGG-----GGACCACTGGACCGGAGCGTGAAGAAGAA	430
Db	125	TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluAspGluAspGlu	144
Qy	431	GACTGGAGACAACCAAGAGAAGAT-----TGGAGGCGACCAAGTCATCAGCAGCCACGG	484
Db	145	AspGluGlnAspGluArgGlnPheProPheProHisGlnLysGluGlu	164
Qy	485	AAATAAAGCCCCGAAGAGAAGAGAAACAAGAGTGGGGAACAACAGGTAGCCATGTG	544

```
Db 165 ArgGlyGlnGluGluAspGluGluGlnGlnArgGluSerGluGluSerGluAsp 184
QY 545 AGGGAA-----GAAACATCTCGAACAACCCCTTTCTACTTCCCGTCAAGCGGTTTAGC 598
Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
QY 599 ACCGGTACGGGAACCAAAACGGTAGATCCGGGTCTCGAGAGTTTGACCAAGGTCA 658
Db 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY 659 AGCAGTTTCAGATCTCCAGATACCCGTATTGTGCAGATCCAGGCCAAACCTACACT 718
Db 225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleGluPheAsnSerLysProAsnThr 244
QY 719 CTGTGTTCTCCCAAGCAGCGTGTATGCTGATAACATCTGTTATTCAGCAAGGCAAGCC 778
Db 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
QY 779 ACCGTGACCGTAGCAATGCAATATAACAGAAAGAGCTTTAATCTTACGAGGCCATGCA 838
Db 265 IleLeuSerLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
QY 839 CTCAGATCCCATCCCGTTTCATTTCTTACATCTTGAACCGCCATCACACCAAGACCTC 898
Db 285 LeuArgValProSerGlyThrThrTyrTyrValValAsnProAspAsnGlnAsnLeu 304
QY 899 AGATGAGTAAATATCTCCATGCCCGTTAAACACACCCCGGCAGTTTCAGAGATTTCTCCCG 958
Db 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPheLeu 324
QY 959 GCAGCAGCGCAGCAACCATCATCTTCTGACGGGTTTCAGCAGGAATACGTTGGAGGCC 1018
Db 325 SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY 1019 GCCTTCAATCGGAATTCATGATGATACCGAGGCTCTGTAGAGAGATGCAGAGGT 1078
Db 345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly--- 363
QY 1079 GAGCAAGAGAGAGAGAGCGGAGCGGATGAGTACTCGGAGTAGTGAGAACATGAGGA 1138
Db 364 -----GlnGlnGlnGlyGluArgLeu-----GlnGluSer 374
QY 1139 GTGATAGTCAAGTGTCAAAGAGAGCGCTGTGAAGAACTTACTAAGCAGCGTAAATCCGTC 1198
Db 375 ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
QY 1199 TCAAGAAAGGCTCCGAAGAAGAGGAGATATCACCAACCAATCACTTGAGAGAAGC 1258
Db 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
QY 1259 GAGCCGATCTTTCTAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAGAGAAGAAC 1318
Db 412 AspProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
QY 1319 CCCAGCTTCAGGACCTGGACATGATGCTCCTCCTGTGTAGATCAAGAGAGGCTTTG 1378
Db 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCCACTTCAACTCAAGGCGATGTTATCGTCTGCTCAACAAAGAACTGGA 1438
Db 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTTGAACCTGTGGCTGTGAAGAAAGAGCAACACAGAGGCGGCGGGGAAGAGAG 1498
Db 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGln 489
QY 1499 GAGGACGAGACCAAGAGGAGGAGTAAGTAACAGAGAGTGCGTAGGTACACAGCGAGG 1558
Db 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
QY 1559 TTGAAGGAGCGGATGTTTCATCATGCCAGAGGCTCATCCAGTAGCTCAACGCTTCC 1618
Db 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
```

---

```
QY 1619 TCCGAATCCATCTGCTTGGCTTCGATCAACGCTGAAAAACACACAGAAATCTTCCTT 1678
Db 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
QY 1679 GCAGGTGATAAGCAATGTGTAGACCAAGAGATAGAGAACGAGCGAGGATTTAGCATTC 1738
Db 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
QY 1739 CTGGGTCCGGTGAACAAGTTGAGAGCTCATCAAAACCCAGAACGAATCTCATTGTG 1798
Db 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
QY 1799 AGTGTCTGCTCAATCTCAATCTCAATCTCGCTGCTCTCTCGAGAAAGAGTCTCTGAG 1858
Db 580 AspAlaGlnProLysLys-----585
QY 1859 AAAGAGGATCAAGAGAGAGGAAAAACCAAGAGGAGGAGGTCCACTCTTTCAATTTTGAAG 1918
Db 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
QY 1919 GCTTTT 1924
Db 603 AlaPhe 604

RESULT 15
US-09-331-631A-25
; Sequence 25, Application US/09331631A
; Patent No. US20020168392A1
; GENERAL INFORMATION:
; APPLICANT: Manners, John M.
; APPLICANT: Marcus, John Paul
; APPLICANT: Goulter, Kenneth C.
; APPLICANT: Green, Jodie L.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; FILE REFERENCE: CULN23.001AEC
; CURRENT APPLICATION NUMBER: US/09/331,631A
; PRIOR FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: PCT/AU97/00874
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: AU PO 4275
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Soybean (Glycine max)
US-09-331-631A-25

Alignment Scores:
Pred. No.: 2,49e-96 Length: 605
Score: 1281.00 Matches: 285
Percent Similarity: 59.21% Conservative: 107
Best Local Similarity: 43.05% Mismatches: 174
Query Match: 35.96% Indels: 96
DB: 9 Gaps: 18

US-10-728-051-1 (1-2032) x US-09-331-631A-25 (1-605)
QY 50 ATGAGAGGAGGAGGTTTCTCCATGTATGCTGTGCTAGGATCTTGTCTCGGCTTCAGTT 109
Db 2 MetaArgAlaArgPhePro-----LeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
QY 110 TCTGCAACGCATGCCCAAGTCATCATCTTACCAGAGAAACAGAGAACCCCTGCGCCAG 169
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGCATCGGAGTCT 229
Db 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGCTGCACCAAGCTCGAGTATGAT-----CCTCCTTGT 262
```

```
Db 57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyGluLeuProArg--- 75
||||| ||||| :|||
QY 263 GTCTATGATCCTCGAGGACACACTGGCACCACCAACCAAGTTCCTCCAGGGAGCGG 322
||||| |||||
Db 76 -----ProArgProArg-----ProGlnHisProGlu 84
|||||
QY 323 ACAGTGGCCGCAACCCGAGAGACTAGATGATGAC----- 358
:||||| :|||
Db 85 ArgGluProGlnGlnProGlyGluLysGluGluAspGluGluAspGluProArgProIle 104
|||||
QY 359 -----CGCGTCAACCCCAAGAGAG-----GAAGGAGGCCGA 391
||||| ||||| :|||
Db 105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGluGlnGlu 124
|||||
QY 392 TGG-----GGACCAGCTGGACCCGAGGAGCGGTGAAGAGAGAA 430
||||| ||||| :|||
Db 125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGlu 144
|||||
QY 431 GACTGGAGACAACAAGAGAGAT-----TGGAGGCGCAACAGTCAATCAGACCCACGG 484
||||| :|||
Db 145 AspGluGlnAspGluArgGlnPheProPheProArgProProHisGlnLysGluGlu 164
|||||
QY 485 AAAATAAGGCCGCAAGAGAGAGAGAAAGAGTGGGGAACACACCGTAGCCATGTG 544
||||| :|||
Db 165 ArgAsnGluGluAspGluAspGluGlnGlnArgGluSerGluGluSerGluAsp 184
|||||
QY 545 AGGGAA-----GAAACATCTCGAACAACCTTTCTACTTCCTCCGTCAGCGGTTTAGC 598
||||| |||||
Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
|||||
QY 599 ACCGGTACGGGAACCAACCGTAGATCCGGTCTCGAGAGGTTTGACCAAGGTCA 658
||||| :|||
Db 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
|||||
QY 659 AGCAGTTTCAGATCTCCAGATCCCGTATTGTCAGATCCAGGCCAACCTACACT 718
||||| :|||
Db 225 ProGlnLeuGlnAsnLeuArgPheTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
|||||
QY 719 CTGTGTTCTCCCAAGCAGCTGATGTGTATAACATCTCTGTTATCCAGCAAGGCGCAAGCC 778
||||| :|||
Db 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
|||||
QY 779 ACCGTGACCTAGCAATGCAATACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCA 838
||||| :|||
Db 265 IleLeuSerLeuValAsnAsnAspAspArgSerTyrArgLeuGlnSerGlyAspAla 284
|||||
QY 839 CTCAGATCCCATCCCGTTTCATTTCTCTACATCTTGACCGCCATGACACCAAGACCTC 898
||||| :|||
Db 285 LeuArgValProSerGlyThrThrTyrTyrValValAsnProAspAsnAsnGluAsnLeu 304
|||||
QY 899 AGAGTAGCTAAATCTCCATGCCGCTTAACACACCCGCGCAGTTTGAGGATTTCTCCCG 958
||||| :|||
Db 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 324
|||||
QY 959 GCGAGCAGCGGACCAATCATCTCTCAGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
||||| :|||
Db 325 SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
|||||
QY 1019 GCCTTCAATCCGGAATTCATAGATACGAGGCGTGTGTAGAGAGATGCGAGGCT 1078
||||| :|||
Db 345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly--- 363
|||||
QY 1079 GAGCAGAGGAGGAGGAGGAGCGGATGAGTACTCGGAGTAGTGAGAACATGAAGGA 1138
|||||
Db 364 -----GlnGlnGlnGlnArgLeu-----GlnGluSer 374
|||||
QY 1139 GTGATGTCAAAGTGTCAAAGGAGCACGTTGAAGAACTTACTAAGCACGCTAAATCCGTC 1198
||||| :|||
Db 375 ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
|||||
QY 1199 TCAAGAAAGGCTCCGAAGAGAGGAGAGATATCCACCAACCAATCACTTGAGAGAGGC 1258
|||||
```

```
Db 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
|||||
QY 1259 GAGCCGATCTTTCTAACTTTTGGAGTATTATTTGAGGTGAAGCCAGACAAGAAC 1318
:||||| |||||
Db 412 AspProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
|||||
QY 1319 CCCAGCTTCAGGACCTGGACATGATGCTCCTCCTGTGTAGATCAAGAGAGAGCTTTG 1378
||||| :|||
Db 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
|||||
QY 1379 ATGCTCCACACTTCAACTCAAAGGCATGGTTATCGTCTGCTCAACAAAGCACTGGA 1438
||||| :|||
Db 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
|||||
QY 1439 AACCTTGAATCTGCTGCTGTAAAGAAAGAGACAACAAGAGGCGCGGGGAAGAG 1498
||||| :|||
Db 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGluGln 489
|||||
QY 1499 GAGGACGACGACGACGAGGAGGAGGAGGAGTAAACAGAGAGGTGCTAGGTACACAGCGAGG 1558
|||||
Db 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
|||||
QY 1559 TTGAAGGAGGCGATGTTTCATGCTCCAGCAGCTCATCCAGTAGCCATCAACGCTTC 1618
||||| :|||
Db 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
|||||
QY 1619 TCCGAATCCATCTGCTTGGCTTCCGTTATCAACGCTGAAACCAACACAGAAATCTTCCTT 1678
||||| :|||
Db 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
|||||
QY 1679 GCGGTGATAGGACAAATGTATGATACCAAGATAGAGAACGACGAGGATTTAGCATTC 1738
||||| :|||
Db 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
|||||
QY 1739 CTTGGCTCGGTGAACAAGTTGAGAAGCTCATCAAAACCCAGAAAGAAATCTCACTTTGTG 1798
||||| :|||
Db 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
|||||
QY 1799 AGTGTCTGCTCAATCTCAATCTCAATCTCCGCTCTCTCTGAGAAAGAGTCTCCTGAG 1858
||||| :|||
Db 580 AspAlaGlnProLysLys----- 585
|||||
QY 1859 AAAGAGATTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1918
||||| :|||
Db 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
|||||
QY 1919 GCTTTT 1924
|||||
Db 603 AlaPhe 604
|||||
```

Search completed: April 8, 2005, 06:48:50  
Job time : 276.254 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 8, 2005, 04:11:27 ; Search time 54.9253 Seconds  
(without alignments)  
5523.384 Million cell updates/sec

Title: US-10-728-051-1

Perfect score: 3562

Sequence: 1 aataacatatataattcatc.....cgttgtggtgtttcttc 2032

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Xgapop 10.0			
Xgapext 0.5			
Ygapop 6.0			
Ygapext 7.0			
Delop 6.0			
Delext 7.0			

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10728051/runat\_07042005\_125046\_14072/app.query.fasta\_1.4757  
-DB=Issued Patents AA -CPM=fastan -SUFFIX=rai -MINMATCH=0 1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10728051@cgn 1 1 122 @runat\_07042005\_125046\_14072 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3278	92.0	626	4	US-09-106-872A-4
2	1700	47.7	335	4	US-09-106-872A-17
3	1279	35.9	605	1	US-07-955-905A-24
4	1267	35.6	571	1	US-07-955-905A-25
5	1068	30.0	410	1	US-07-955-905A-26
6	811	22.8	566	1	US-07-955-905A-22
7	811	22.8	566	1	US-07-955-905A-22
8	788	22.1	421	1	US-07-955-905A-27
9	777	21.8	587	1	US-07-955-905A-23
10	591.5	16.6	489	4	US-09-424-283-3
11	569	16.0	524	4	US-09-424-283-1
12	559.5	15.7	448	4	US-09-323-195A-18

13	514.5	14.4	409	4	US-09-424-283-4	Sequence 4, Appli
14	513.5	14.4	523	4	US-09-323-195A-17	Sequence 17, Appl
15	501	14.1	444	4	US-09-424-283-2	Sequence 2, Appli
16	183	5.1	762	4	US-09-352-991A-29423	Sequence 29423, A
17	173	4.9	1564	4	US-10-144-198-2	Sequence 2, Appli
18	173	4.9	1564	4	US-10-144-198-4	Sequence 4, Appli
19	166.5	4.7	720	4	US-09-352-991A-21881	Sequence 21881, A
20	164.5	4.6	1266	4	US-09-352-991A-30851	Sequence 30851, A
21	163	4.6	714	4	US-09-949-016-10296	Sequence 10296, A
22	160	4.5	630	4	US-09-352-991A-26324	Sequence 26324, A
23	160	4.5	801	4	US-09-352-991A-29274	Sequence 29274, A
24	160	4.5	1898	1	US-08-056-200-94	Sequence 94, Appl
25	160	4.5	1898	2	US-08-056-200-94	Sequence 94, Appl
26	160	4.5	1898	4	US-09-538-092-1280	Sequence 1280, Ap
27	159	4.5	1043	4	US-09-352-991A-28885	Sequence 28885, A
28	158.5	4.4	1075	4	US-09-352-991A-18387	Sequence 18387, A
29	157	4.4	1780	1	US-08-769-309A-5	Sequence 5, Appli
30	157	4.4	1780	3	US-08-994-570-5	Sequence 5, Appli
31	157	4.4	1781	4	US-09-361-403-13	Sequence 13, Appl
32	157	4.4	2079	4	US-09-949-016-8301	Sequence 8301, Ap
33	155	4.4	754	4	US-09-976-594-375	Sequence 375, App
34	155	4.4	760	4	US-09-949-016-11129	Sequence 11129, A
35	155	4.4	863	4	US-09-352-991A-26099	Sequence 26099, A
36	154	4.3	650	4	US-09-352-991A-20375	Sequence 20375, A
37	153.5	4.3	1233	4	US-09-688-188B-89	Sequence 89, Appl
38	153.5	4.3	1233	4	US-09-291-417D-89	Sequence 89, Appl
39	152.5	4.3	580	4	US-09-248-796A-14473	Sequence 14473, A
40	152	4.3	1162	2	US-08-728-323A-2	Sequence 2, Appli
41	152	4.3	1162	3	US-09-298-568-2	Sequence 2, Appli
42	152	4.3	1162	4	US-09-410-399-2	Sequence 2, Appli
43	152	4.3	1162	4	US-09-894-273-2	Sequence 2, Appli
44	152	4.3	1427	4	US-09-252-991A-27005	Sequence 27005, A
45	151.5	4.3	917	4	US-09-252-991A-25101	Sequence 25101, A

#### ALIGNMENTS

RESULT 1

US-09-106-872A-4

; Sequence 4, Application US/09106872A

; Patent No. 6486311

; GENERAL INFORMATION:

; APPLICANT: Burks Jr., A. Wesley

; APPLICANT: Stanley, J. Steven

; APPLICANT: Cockrell, Gael

; APPLICANT: King, Nina E.

; APPLICANT: Sampson, Hugh A.

; APPLICANT: Helm, Ricki M.

; APPLICANT: Bannon, Gary A.

; TITLE OF INVENTION: Peanut Allergens and Methods

; FILE REFERENCE: HS 103 CIP

; CURRENT APPLICATION NUMBER: US/09/106,872A

; CURRENT FILING DATE: 1999-06-29

; PRIOR APPLICATION NUMBER: PCT/US96/15222

; PRIOR FILING DATE: 1996-09-23

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 626

; TYPE: PRT

; ORGANISM: Arachis hypogaea

; FEATURE:

; OTHER INFORMATION: Amino Acids 25-34 are Ara H 1 binding epitope,

; OTHER INFORMATION: peptide 1

; OTHER INFORMATION: Amino Acids 48-57 are Ara H 1 binding epitope,

; OTHER INFORMATION: peptide 2

; OTHER INFORMATION: Amino Acids 65-74 are Ara H 1 binding epitope,

; OTHER INFORMATION: peptide 3

; OTHER INFORMATION: Amino Acids 89-98 are Ara H 1 binding epitope,

; OTHER INFORMATION: peptide 4

; OTHER INFORMATION: Amino Acids 97-106 are Ara H 1 binding epitope,

; OTHER INFORMATION: peptide 5

; OTHER INFORMATION: Amino Acids 107-116 are Ara H 1 binding epitope,

OTHER INFORMATION: peptide 6  
OTHER INFORMATION: Amino Acids 123-132 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 7  
OTHER INFORMATION: Amino Acids 134-143 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 8  
OTHER INFORMATION: Amino Acids 143-152 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 9  
OTHER INFORMATION: Amino Acids 294-303 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 10  
OTHER INFORMATION: Amino Acids 311-320 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 11  
OTHER INFORMATION: Amino Acids 325-334 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 12  
OTHER INFORMATION: Amino Acids 344-353 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 13  
OTHER INFORMATION: Amino Acids 393-402 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 14  
OTHER INFORMATION: Amino Acids 409-418 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 15  
OTHER INFORMATION: Amino Acids 461-470 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 16  
OTHER INFORMATION: Amino Acids 498-507 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 17  
OTHER INFORMATION: Amino Acids 525-534 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 18  
OTHER INFORMATION: Amino acids 539-548 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 19  
OTHER INFORMATION: Amino acids 551-560 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 20  
OTHER INFORMATION: Amino acids 559-568 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 21  
OTHER INFORMATION: Amino acids 578-587 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 22  
OTHER INFORMATION: Amino acids 597-606 are Ara H 1 binding epitope,  
OTHER INFORMATION: peptide 23

US-09-106-872A-4

Alignment Scores:  
Pred. No.: 1,51e-308 Length: 626  
Score: 3278.00 Matches: 625  
Percent Similarity: 99.84% Conservative: 0  
Best Local Similarity: 99.84% Mismatches: 1  
Query Match: 92.03% Indels: 0  
DB: 4 Gaps: 0

US-10-728-051-1 (1-2032) x US-09-106-872A-4 (1-626)

QY 50 ATGAGGGAGGGTTTCTCCACTGATGCTGTTCTAGGATCCTGTCTGGCTTCAGTT 109  
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20  
QY 110 TCTGCAACGATGCCAAGTCCATCACCTTACCAGAAGAAACAGAAACCCCTCGGCCAG 169  
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40  
QY 170 AGTGCCCTCCAGATGTTCTCAACAGGAACCGGATCACTTGAAGCAAAAGGCATCGAGTCT 229  
DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60  
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTGTGTCTATGATCCTCGAGGACACACTGGC 289  
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80  
QY 290 ACCACCAACCAAGCTTCCCTCCAGGGGAGCGGACAGCTGGCCGCCAACCCCGGAGACTAC 349  
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100  
QY 350 GATGATACCCCTCAACCCCGAAGAGAGAGAGAGCGCGATGGGACCCAGCTGGACCG 409  
DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTyrGlyProAlaGlyPro 120  
QY 410 AGGAGCGTGAAGAGAGAGAGACTGGAGACACCAAGAGAGATTCGAGGCCCAAGT 469

DB 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140  
QY 470 CATCAGCAGCCAGCAAAATAAGCCGCAAGAGAGAGAGAAACAAGAGTGGGCAACA 529  
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160  
QY 530 CCAGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTTTCTACTCCCTCAAGG 589  
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180  
QY 590 CGGTTTAGCCCGCTACGGNACCAAAACCGTAGGATCCGGTCTCCAGAGGTTTTCAC 649  
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200  
QY 650 CAAAGGTCAAGGCAGTTTCCAGAACTCCAGAAATCACCGTATTGTGTGTCAGATCGAGCCAAA 709  
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220  
QY 710 CTTAACACTCTTGTCTTCCCAAGCACGCTGATGCTGATAACATCCTTGTGTATCCAGCA 769  
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240  
QY 770 GGGCAAGCCCGCTGACCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTCACCGAG 829  
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260  
QY 830 GGCCATGCACCTCAGAAATCCCATCCCGTTTCAATTCCTACATCTTGAACCGCATGACAA 889  
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280  
QY 890 CAGAACCTCAGAGTAGCTAAATCTCATGCCCCGTAAACACACCCGGCCAGTTTTCAGGAT 949  
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300  
QY 950 TTCTTCCCGCGCAGCAGCCGAGACCAATCATCTACTTGCAGGGCTTCAGCAGGAAATAG 1009  
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGluPheSerArgAsnThr 320  
QY 1010 TTGGAGGCGCCTTCAATGCGGAATTCATAGATACGAGGGTCTGTGTAGACAGAGAT 1069  
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340  
QY 1070 GCAGAGGTGACCAAGAGAGAGAGAGCGGAGCCATGGAGTACTCGGAGTAGTCAGAC 1129  
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgTyrTrpSerThrArgSerSerGluAsn 360  
QY 1130 AATGAAGGAGTGCATAGTCAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTTAAGCACGCT 1189  
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380  
QY 1190 AAATCCGCTCTCAAAGAAAGGCTCCGAAAGAGGAGATATCAACCAACCAATCAACTTG 1249  
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400  
QY 1250 AGAAGAGCGAGCCCGATCTTCTTAACACTTGGAGTATTTTTCAGGTGAGCCAGAC 1309  
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420  
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTCGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369  
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440  
QY 1370 GGAGCTTTGATGCTCCCACTTCAACTCAAGGCCATGGTTATCTGCTCGTCAACAA 1429  
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValValVal 460  
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAAGAAAGAGCAACAACAGAGGCGAGCCGG 1489  
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480  
QY 1490 GAAGAGAGGAGGAGCAGCAAGCAAGAGAGGAGGAGGAGTAACAGAGAGGTGCTAGGTAC 1549  
DB 481 GluGluGluGluAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 500

QY	1550	ACAGCAGGTTGAAGGAAGCGATGCTTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
Db	501	ThrAlaArgLeuIysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AAGCCTTCCTCCGNACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACACCAACAGA	1669
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCCTCCAGTGATAAGGCAATGTGATAGACCAGATAGAGAAGCAAGCGAAGGAT	1729
Db	541	IlePheLeuAlaGlyAspIysAspAsnValIleAspGlnIleGluIysGlnAlaLysAsp	560
QY	1730	TTAGCATTCCTCGGTCGGGTGAACAGTTGTGAGAAGCTCATCAAAACCAAGAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluIysLeuIleIysAsnGlnIysGluSer	580
QY	1790	CACTTTGTGAGTGCTCGCTCCTCAATCTCAATCTCAATCTCCGTCGTCTCCTGAGAAAGAG	1849
Db	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluIysGlu	600
QY	1850	TCTCCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGAGGGAAGGTCACCTCCTTTTCA	1909
Db	601	SerProGluIysGluAspGlnGluGluAsnGlnGlyGlyIysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGCCTTTTAAC	1927
Db	621	IleLeuIysAlaPheAsn	626

## RESULT 2

US-09-106-872A-17  
; Sequence 17, Application US/09106872A

Patent No. 6486311

```

1  / GENERAL INFORMATION:
2  / APPLICANT: Burks Jr., A. Wesley
3  / APPLICANT: Stanley, J. Steven
4  / APPLICANT: Cockrell, Gael
5  / APPLICANT: King, Nina E.
6  / APPLICANT: Sampson, Hugh A.
7  / APPLICANT: Helm, Ricki M.
8  / APPLICANT: Bannon, Gary A.
9  / TITLE OF INVENTION: Peanut Allergens and Methods
10 /
11 / FILE REFERENCE: HS 103 CIP
12 / CURRENT APPLICATION NUMBER: US/09/106, 872A
13 / CURRENT FILING DATE: 1999-05-29
14 / PRIOR APPLICATION NUMBER: PCT/US96/15222
15 / PRIOR FILING DATE: 1996-09-23
16 / NUMBER OF SEQ ID NOS: 23
17 / SOFTWARE: Patent In Ver. 2.1

```

Alignment Scores:	8.73e-156	Length:	335
Pred. No.:	1700.00	Matches:	335
Score:	99.41%	Conservative:	0
Percent Similarity:	99.41%	Mismatches:	0
Best Local Similarity:	99.41%	Indels:	2
Query Match:	47.73%	Gaps:	1
DB:	4		

US-10-728-051-1 (1-2032) x US-09-106-872A-17 (1-335)

917	ATGCCGTTTAAACACACCCGCCAGTTTGGAGATTCTTCCGGCGAGCAGCCGAGACCAA	976
Qy		
1	MetProValAsnThrProGlyGlnPheGluAspPheProLaSerSerArgAspGln	20
Db		
977	TCATCTTACTTGCGGGCTTCAGCAGAAATACGTTGGAGGCCGCTTCAATCGGGAAATC	1036
Qy		
21	SerSerTyrLeuGlnGlyPheSerArgAsnThrLeuGluAlaAlaPheAsnAlaGluPhe	40
Db		

QY	1037	AATGAGATACGGAGGCTGCTGTTAGAAAGAAATGCAGGAGGTGACCAAGAGAGAGAGGG	1098
DB	41	AsnGluIleArgValLeuLeuGluGluAsnAlaGlyGlyGluGluGluArgGly	60
QY	1097	CAGAGGCGATGGAGTACTCGAGCTAGTGAACAATACTGAAGGAGTGTAGTCAAAAGTCTCA	1156
DB	61	GlnArgArgTrpSerThrArgSerSerGluAsnAsnGluGlyValIleValSer	80
QY	1157	AAGGAGCACGTTGAAGAAGCTTACTTAAGCAGCGCTAAATCCGTCTCAAGAAGAGCTCCGAA	1216
DB	81	LysGluHisValGluGluLeuThrLysHisAlaLysSerValSerLysGlySerGlu	100
QY	1217	GAAGAGGAGATATCACCAACCACTCAACTTGAAGAGGCGAGCCGATCTTTCTTAAC	1276
DB	101	GluGluGlyAspIleThrAsnProIleAsnLeuArgGluGluProAspLeuSerAsn	120
QY	1277	AACCTTTGGGAAGTTATTTCAGGGTGAAGCCAGACAAGAAGAACCCAGCTTCAGGACCTG	1336
DB	121	AsnPheGlyLysLeuPheGluValLysProAspLysLysAsnProGlnLeuGlnAspLeu	140
QY	1337	GACATGATGCTACCTGTGTAGAGATCAAGAAGAGGAGCTTTGTAGTCTCCACACTTCAAC	1396
DB	141	AspMetMetLeuThrCysValGluIleLysGluValAlaLeuMetLeuProHisPheAsn	160
QY	1397	TCAAAGCCCATGTTATCGTCGTCTCAACAAGAAGAACTGTAACCTTGAACCTCGTGGCT	1456
DB	161	SerLysAlaMetValIleValValValAsnLysGlyThrGlyAsnLeuGluLeuValAla	180
QY	1457	GTAAGAAAAGAGCAACAACAGAGGGGACGCGGAAGAGAGAGGAGGACGAAGACGAGAA	1516
DB	181	ValArgLysGluGlnGlnArgGlyArgArgGluGluGluGluGluAspGluGluGlu	200
QY	1517	GAGGAGGGAAGTAAACAGAGAGGTCGCTAGGTACACACGACGAGTTGAAGGAAGCGCATGTG	1576
DB	201	GluGluGlySerAsnArgGluValArgArgTyrThrAlaArgLeuLysGluGlyAspVal	220
QY	1577	TTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTT	1636
DB	221	PheIleMetProAlaAlaHisProValAlaIleAsnAlaSerSerGluLeuHisLeuLeu	240
QY	1637	GGCTTCGGTATCAACCTGAAACACACCAACAGAACTTCTCTCCAGGTGTATAGGACAAT	1696
DB	241	GlyPheGlyIleAsnAlaGluAsnAsnHisArgIlePheLeuAlaGlyAspLysAspAsn	260
QY	1697	GTGATAGACAGATAGAGAAGCAAGCGAAGATTTAGCATTCCTCGGTGCGGTGAACAA	1756
DB	261	ValIleAspGlnIleGluLysGlnAlaLysAspLeuAlaPheProGlySerGlyGluGln	280
QY	1757	GTTGAGAAGCTCATCAAAACCAAGAAGGAATCTCATTTGTGTAGTGCTCGCTCTCAATCT	1816
DB	281	ValGluLysLeuIleLysAsnGlnLysGluSerHisPheValSerAla-----GlnSer	298
QY	1817	CAATCTCAATCTCCGTCGTCTCTCTGAGAAGAGTCTCTGAGAAGAGGATCAAGAGGAG	1876
DB	299	GlnSerGlnSerProSerProGluLysGluSerProGluLysGluAspGlnGluGlu	318
QY	1877	GAATAACCAAGGAGGAGGCTCCACTCTTCAATTTTGAAGGCTTTTAAC	1927
DB	319	GluAsnGlnGlyLysGlyProLeuLeuSerIleLeuLysAlaPheAsn	335
RESULT 3			
US-07-955-905A-24			
; Sequence 24, Application US/07955905A			
; Patent No. 5770433			
; GENERAL INFORMATION:			
; APPLICANT:			
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND			
; PRECURSOR			
; NUMBER OF SEQUENCES: 28			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			











```
QY 1298 GTGAAGCCACAGAGAACCCAGCTTCAGGACCTGAGATGATCTCACTGTGTA 1357
Db 369 AlaCysProGluAspPheSer---GlnPheGlnAsnMetAspValAlaValSerAlaPhe 387
QY 1358 GAGATCAAAAGAGAGAGCTTGTATGCTCCACACACTTCAACTCAAGGCCATGTTATCGTC 1417
Db 388 LysLeuAsnGlnGlyAlaIlePheValProHisTyrAsnSerLysAlaThrPheValVal 407
QY 1418 GTCGTCAACAAAGAACTGGAACCTTGAACCTGCTGCTGTATAGAAAAGAGCAAA--- 1474
Db 408 PheValThrAspGlyTyrGlyTyrAlaGlnMetAlaCysProHisLysLeuSerArgGlnSer 427
QY 1475 -----CAGAGGGAGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1525
Db 428 GlnGlySerGlnSerGlyArgGlnGlnValLysAlaPheValPheValAla 447
QY 1526 AGTAACAGAGAGGTGCTAGGTATACACAGCGAGGTGTAAGAGAGCGATGTTTCATCATG 1585
Db 448 ThrPheGlyGluPheGlnGlnValLysAlaPheValPheValAla 467
QY 1586 CCAGCAGCTCATCAGTAGCCATCAAGCTTCTCCGAA-----CTCCATCTGCTGGC 1639
Db 468 ProAlaGlyHisAlaValThrPhePheAlaSerLysAspGlnProLeuAsnAlaValAla 487
QY 1640 TTCGTATCAACGCTGAAACAAACACAGAACTTCTCTTCAGGTGATAGAGCAATGTG 1699
Db 488 PheGlyLeuAsnAlaGlnAsnAsnGlnArgIlePheLeuAlaGly---LysLysAsnLeu 506
QY 1700 ATAGACCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1759
Db 507 ValArgGlnMetAspSerGluAlaLysGluLeuSerPheGlyValProSerLysLeuVal 526
QY 1760 GAGAAGCTCATCAAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1819
Db 527 AspAsnIlePheAsnAsnProAspGluSerTyrPheMetSerPheSerGlnGlnArgGln 546
QY 1820 TCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGAGAG 1879
Db 547 -----ArgArgAspGlu 550

RESULT 7
US-07-955-905A-22
; Sequence 22, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Theobroma cacao
; FEATURE:
; NAME/KEY: Protein
```

```
; LOCATION: 1..566
; OTHER INFORMATION: /note= "67 kD Precursor Protein"
US-07-955-905A-22
```

```
Alignment Scores: 1.75e-69 Length: 566
Pred. No.: 811.00 Matches: 195
Score: 50.95% Conservative: 127
Percent Similarity: 30.85% Mismatches: 218
Best Local Similarity: 22.77% Indels: 92
Query Match: 1 Gaps: 19
DB: 1
```

US-10-728-051-1 (1-2032) x US-07-955-905A-22 (1-566)

```
QY 65 TCTCCACTGATGCTGCTGCTAGGATCCTTCTCTGCTTCAGTTTCTGCAACCATGCC 124
Db 6 SerProPheIleValLeuIlePheSerLeuLeuLeuSerPheAlaLeuLeuCysSerGly 25
QY 125 AAGTCATCACCTTACAGAGAAACAGAGAACCCCTGCGCCAG-----AGTGTC 175
Db 26 ValSerAlaTyrGlyArgLysGlnTyrGluArgAspProArgGlnGlnTyrGluGlnCys 45
QY 176 CTCCAGAGTGTCAACAGGAA---CCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGC 232
Db 46 GlnArgArgCysGluSerGluAlaThrGluGluArgGluGlnGlnCysGluGlnArg 65
QY 233 TGCACCAAGCTCGAGTATGATCCTCGTGTGTCTATGATCTCTCGAGGACACACTGGCACC 292
Db 66 CysGluArg---GluTyrLysGluGln----- 73
QY 293 ACCAACCAACGTTCCCTCCAGGGAGCGGACACGTCGCCGCCCAACCCGGAGACTACGAT 352
Db 74 -----GlnArgGlnGlnGluGluLeuGlnArgGlnTyrGln----- 86
QY 353 GATGACCGCGCTCAACCCCGAAGAGAGAGAGAGAGCGCCGATGGGGACCAGCTGGACCGAGG 412
Db 87 -----GlnCysGlnGlyArgCysGlnGlnGlnGln-----GlyGlnArg 100
QY 413 GAGCGTGAA-----AGAGAAGAGACTGGAGACAACAAGAGAGAGATGGAGG----- 460
Db 101 GluGlnGlnGlnCysGlnArgLysCysTrpGluGlnTyrLysGluGlnGluGlnArgGlyGlu 120
QY 461 CGACCAAGTCATCAGCAGCCAGGAAATTAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAG 520
Db 121 HisGluAsnTyrHisAsnHisLysLysAsnArgSerGluGluGluGluGlnGln--- 139
QY 521 TGGGGAAACACAGGTAGCCATCTGAGGGAGAGAAACATCTCGGAAACAACCTTTCTACTTC 580
Db 140 -----ArgAsnAsnProTyrTyrPhe 146
QY 581 CGCTCAAGCGG---TTTAGCACCCGCTACGGGAACCAACCGTAGGATCCGGTCTCTG 637
Db 147 ProLysArgArgSerPheGlnThrArgPheArgAspGluGluGluGluGluGluGluLeu 166
QY 638 CAGAGGTTTGACCAAGGTCAAGGCGAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGACG 697
Db 167 GlnArgPheAlaGluAsnSerProProLeuLysGlyIleAsnAspTyrArgLeuAlaMet 186
QY 698 ATCAGGCCCAACCTAACACTCTGTTCTTCCAGACGCTGATGCTCTATACATCCTT 757
Db 187 PheGluAlaAsnProAsnThrPheIleLeuProHisHisCysAspAlaGluAlaIleTyr 206
QY 758 GTTATCCAGCAAGGCAAGCCCGTAGCCAGTAAATGGCAATTAACAGAAAGAGCTTT 817
Db 207 PheValThrAsnGlyLysGlyThrIleThrPheValThrHisGluAsnLysGluSerTyr 226
QY 818 AATCTTTGACGAGGCCCATGCATCTCAGATCCCATCCGTTTCTATTTCTTACATCTTGAAC 877
Db 227 AsnValGlnArgGlyThrValValSerValProAlaGlySerThrValTyrValValSer 246
QY 878 CGCCATGACAAACAGAACCTCAGATAGTAAATCTCCATGCCCTTAAACACACCCGGC 937
Db 247 GlnAspAsnGlnLysLeuThrIleAlaValLeuAlaLeuProValAsnSerProGly 266
```

```
QY 938 CAGTTTGAAGATTCTTCCCGGCGAGCAGCGAGCAACATCACTCTACTTTCGAGGGCTTC 997
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
Db 267 LysTyrGluLeuPhePheProAlaGlyAsnAsnLysProGluSerTyrTyrGlyAlaPhe 286
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 998 AGCAGGAATACGTTGAGGCGCTTCAATCCGGAATCAATGATACGAGGGTCTG 1057
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 287 SerTyrGluValLeuGluThrValPheAsnThrGlnArgGluLysLeuGluGluLeu 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1058 TTAGAAGAAATGACGAGGTGTAGCAAGAGAGAGAGAGGCGAGCGCATGGAGTACTCGG 1117
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
Db 307 -----GluGluGlnArgGlyGlnLysArg----- 314
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1118 AGTAGTGAGAACATGAAGAGTGATAGTCAAGTGTCAAGAGACGTTGTAAGAACTT 1177
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
Db 315 ---GlnGlnGlyGlnGlnGlyMetPheArgLysAlaLysProGluGlnIleArgAlaIle 333
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1178 ACTAAGCACGCTAAATCCGCTCAAGAGAGAGAGAGAGGCGAGCGCATGGAGTACTCGG 1237
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
Db 334 SerGlnGlnAlaThrSerProArgHisArgGlyGlyGluArgLeuAla----- 349
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1238 CCAATCAACTTGAAGAGCGAGCGCGATCTTTCTAACCACTTTTGGGAAGTTATTGAG 1297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 ---IleAsnLeuSerGlnSerProValTyrSerAsnGlnAsnGlyArgPhePheGlu 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1298 GTGAAGCCAGAACAGAAACCCCGCTTCAGACCTCGACATGATGCTCACCCTGTGTA 1357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 AlaCysProGluAspPheSer---GlnPheGlnAsnMetAspValAlaValSerAlaPhe 387
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1358 GAGATCAAGAGGAGCTTGATGCTCCACACTTCAACTCAAGGCGCATGGTTATCGTC 1417
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
Db 388 LysLeuAsnGlnGlyAlaIlePheValProHisTyrAsnSerLysAlaThrPheValVal 407
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1418 GTCTCAACAAAGAACTGGAACCTTCAACTCTGCTGTGAAGAAAAGCAACAA--- 1474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 PheValThrAspGlyTyrGlyTyrAlaGlnMetAlaCysProHisLeuSerArgGlnSer 427
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1475 -----CAGAGGCGAGCGCGGGAAGAGAGAGCAGACGAGAGCAAGAGAGGAGGA 1525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 428 GlnGlySerGlnSerGlyArgGlnAspArgArgGlnGlnGluGluSerGluGluGlu 447
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1526 AGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTGTAAGAGCGCATGTTTCATCATG 1585
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
Db 448 ThrPheGlyGluPheGlnGlnValLysAlaProLeuSerProGlyAspValPheValAla 467
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAA-----CTCCATCTGCTGGC 1639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 468 ProAlaGlyHisAlaValThrPhePheAlaSerLysAspGlnProLeuAsnAlaValAla 487
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1640 TTCGGTATCAACGCTGAAACCAACACAGAAATCTTCTTCAGGTGTATAAGCAATGTG 1699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 488 PheGlyLeuAsnAlaGlnAsnAsnGlnArgIlePheLeuAlaGly---LysLysAsnLeu 506
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1700 ATAGACCAGATAGAGAGCAAGCGAAGGATTAGCATTCCTCGGTGCGGTGTAACAAGTT 1759
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
Db 507 ValArgGlnMetAspSerGluAlaLysGluLeuSerPheGlyValProSerLysLeuVal 526
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1760 GAGAAGCTCATCAAAAACAGAGGAATCTCACTTTGTGAGTGTCTGCTCCTCAATCTCAA 1819
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
Db 527 AspAsnIlePheAsnAsnProAspGluSerTyrPheMetSerPheSerGlnGlnArgGln 546
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1820 TCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAAGAGGAGAA 1879
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
Db 547 -----ArgArgAspGlu 550
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
QY 1880 AACCAAGGAGGAGGTCCACTCTCTTCAATTTTG 1915
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
Db 551 ArgArgGly-----AsnProLeuAlaSerIleLeu 560
    : : : : : ||| ||| ||| ||| ||| : : : ||| ||| ||| ||| |||
```

RESULT 8

US-07-955-905A-27

; Sequence 27, Application US/07955905A

; Patent No. 5770433

```
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 421 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Phaseolus vulgaris
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..421
; OTHER INFORMATION: /note= "Vicilin from P. vulgaris"
; US-07-955-905A-27
```

## Alignment Scores:

Pred. No.:	2,466-67	Length:	421
Score:	789.00	Matches:	182
Percent Similarity:	58.44%	Conservative:	88
Best Local Similarity:	39.39%	Mismatches:	106
Query Match:	22.12%	Indels:	86
DB:	1	Gaps:	14

US-10-728-051-1 (1-2032) x US-07-955-905A-27 (1-421)

QY	548	GAAGAACAATCTCGGAACAACCTCTTCTACTTCCCGTCA---AGCGGTTTATGACCCGC	604
Db	30	GluGluGluSerGlnAsnProPheTyrPheAsnSerAspAsnSerTrpAsnThrLeu	49
QY	605	TACGGGAACCAAAACCGTAGGATCCGGGTCTCTCAGAGGTTTACCAAGGTCNAGCAG	664
Db	50	PheLysAsnGlnTyrGlyHisIleArgValLeuGlnArgPheAspGlnGlnSerLysArg	69
QY	665	TTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCTT	724
Db	70	LeuGlnAsnLeuGluAspTyrArgLeuValGluPheArgSerLysProGluThrLeuLeu	89
QY	725	CTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTATCCAGCAAGGGCAAGCCGCTG	784
Db	90	LeuProGlnGlnAlaAspAlaGluLeuLeuValValArgSerGlySerAlaIleLeu	109
QY	785	ACGTAGCAATGGCAATACAGAAAGAGC---TTTAATCTTCAGAGGCCCATGCACTC	841
Db	110	ValLeuValLysProAspArgArgGluTyrPheLeuThrSerAspAsnProIle	129
QY	842	-----AGATCCCATCCGGTTTTCATTCTTCTTACATCTTGAACCGCATGAC	886
Db	130	PheSerAspHisGlnLysIleProAlaGlyThrIlePheTyrLeuValAsnProAspPro	149
QY	887	AACGAGAACCTCAGAGTAGCTAAATCTCCATCCCGCTTAACACACCCCGCCAGTTTGAG	946
Db	150	LysGluAspLeuArgIleIleGlnLeuAlaMetProValAsnAsnPro---GlnIleHis	168
QY	947	GATTCTTCCCGGCGAGCAGCAGACCAATCATCTTCTTCTGAGGGCTTCAGCAGGAAT	1006
Db	169	GluPhePheLeuSerSerThrGluAlaGlnGlnSerTyrLeuGlnGluPheSerLysHis	188
QY	1007	ACGTTGAGGCGCCCTTCAATCGGAATTCGAATGATACGAGGGTGTCTGTAGAGAG	1066
Db	189	IleLeuGluAlaSerPheAsnSerLysPheGluGluIleAsnArgValLeuPheGluGlu	208

```
QY 1067 AATGCAGAGGTGAGCAAGAGGAGAGGCGCAGAGCGATGGAGTACTCGGAGTAGTGA 1126
Db 209 Glu-----GlyGln----- 212
QY 1127 AACAAATGAGAGGTGATGATCAAGAGTGTCAAGAGGAGCAGCTTGAAGAACTTACTAAGCAC 1186
Db 213 -----GluGlyValIleValAsnIleAaspSerGluGlnIleLysGluLeuSerLysHis 230
QY 1187 GCTAAATCCCTCTCAAGAAAGGCTCGAAGAGAGGAGATATCAACCAACCAATCAAC 1246
Db 231 AlalysSerSerSerArgLys-----Ser 238
QY 1247 TTGAGAGAGCGAGCGCCGATCTTTCTAACAACTTTGGGAAGTATTGAGGTGAAGCCA 1306
Db 239 LeuSerLysGlnAaspAsnThrIleGlyAsnGluPheGlyAsnLeuThrGlu---ArgThr 257
QY 1307 GACAAAGAAACCCACAGCTTCAGGACCTGGACATGATCTCACCTGTGTAGAGATCAAA 1366
Db 258 Asp-----AenSerLeuAsnValLeuIleSerSerIleGluMetGlu 271
QY 1367 GAGAGGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCTCGTCAAC 1426
Db 272 GluGlyAlaLeuPheValProHisTyrSerLysAlaIleValIleLeuValAlaAsn 291
QY 1427 AAGGAACTGGAAACCTTGAACCTCGTGGCTGTAAAGAAAGCAACAAAGAGGCGGCG 1486
Db 292 GluGlyGluAlaHisValGluLeuValGly-----ProLysGlyAsn 305
QY 1487 CGGAAGAAGAGGAGGAGCAGAGCAAGAGAGAGAGAGGAGGAAGTAAACAGAGAGGTGCGTAGG 1546
Db 306 LysGluThrLeuGluTyrGluSer----- 313
QY 1547 TACACAGCGAGGTGAAGGAGGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCC 1606
Db 314 TyrArgAlaGluLeuSerLysAspValPheValIleProAlaIleTyrProValAla 333
QY 1607 ATCAACGCTTCTCCGAACCTCACTCTGCTGGCTTCAGCTATCAACGCTGAAACCAAC 1666
Db 334 IleLysAlaThrSerAsnValAsnPheThrGlyPheGlyIleAsnAlaAsnAsnAsn 353
QY 1667 AGAATCTCTCTGAGGTGATAGAGCAATGTGTATAGACAGATAGAGAAG-----CAA 1720
Db 354 ArgAsnLeuLeuAlaGlyLysThrAspAsnValIleSerSerIleGlyArgAlaLeuAsp 373
QY 1721 CGGAGGAT-----TTAGCATTCCTGGTGGTGAACAAAGTTGAGAGCTCATC 1771
Db 374 GlyLysAspValLeuGlyLeuThrPheSerGlySerGlyAspGluValMetLysLeuIle 393
QY 1772 AAAAACCAAGAGGAATCTCACTTTGTGAGTGTCTCTCTCAATCTCAATCTCAATCTCCG 1831
Db 394 AsnLysGlnSerGlySerTyrPheValAspAlaHis----- 405
QY 1832 TCGTCTCTGAGAAAGAGTCTCTGAGAAAGAGATCAAGAGGAGGAACCAAGAGGG 1891
Db 406 -----HisIleGlnGlnGlnLysGlyArg 415
QY 1892 AAGGT 1897
Db 416 LysGly 417
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/955,905A
FILING DATE: 21-JAN-1993
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 587 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Gossypium hirsutum
FEATURE:
NAME/KEY: Protein
LOCATION: 1..587
OTHER INFORMATION: /note= "vicilin from G. hirsutum"
US-07-955-905A-23

Alignment Scores:
Pred. No.: 3,53e-66 Length: 587
Score: 777.00 Matches: 190
Percent Similarity: 51.36% Conservative: 112
Best Local Similarity: 32.31% Mismatches: 208
Query Match: 21.81% Indels: 78
DB: 1 Gaps: 15

US-10-728-051-1 (1-2032) x US-07-955-905A-23 (1-587)
QY 104 TCAGTTTTCGCAACCGATGCCAAGTCATCACCATTACCAAGAGAAACAGAGAACCCCTGC 163
Db 64 SerCysLysSerGlnTyrGlyGluLysAspGlnGlnArgHisArgProGluAspPro--- 82
QY 164 GCCCAGAGTGCTCCAGAGTTGTCAACAGAA-----CCGATGACTTTGAAGCAA 214
Db 83 ---GlnArgTyrGluGluCysGlnGlnGluCysArgGlnGlnGluGluArgGlnArg 101
QY 215 AAGGCATGCGAGTCTGCTGCAACCAAGTCGAGTATGATCTCTGTTGTCTATGATCT 274
Db 102 ProGlnCysGlnArgCysIleLys----- 110
QY 275 CGAGGACACATGGCACCACCAACCAAGTTCCCTCCAGGGGAGCGGACAGTGGCCGC 334
Db 111 -----ArgPheGluGlnGlnGlnGlnSerGlnArg 121
QY 335 CAACCGGAGACTTACGATGATGAC---CGCCGCTCAACCCCGAAGAGAGAGAGCGCGA 391
Db 122 GlnPheGlnGluCysGlnGlnHisCysHisGlnGlnGluGlnArgProGluArgLysGln 141
QY 392 TGGGGACCAAGTGGACCGGAGGAGCTGAAGAGAGAGAGAGAGTGGAGACCAACAGAGAA 451
Db 142 GlnCysValAlaGluCysArgGluArgTyrGlnGluAsnProTyrArgArgGluArgGlu 161
QY 452 GATTGGAGCGCAACAGTATCATCAGCAGCCACCGMAAATAAGGCCCGCAAGAGAGAGGA 511
Db 162 Glu-----GluAlaGluGlu 167
QY 512 GAACAAGAGTGGGAAACACAGATAGCCATGTAGGAGAAAGAAATCTCGGAACACCCCT 571
Db 168 GluThrGluGluGly-----GluGlnGluGlnSerHisAsnPro 180
QY 572 TTCTACTTCCGTCACGCGGTTTACACCCGCTACCGGAACCAACCGTAGGATCCGG 631
Db 181 PheHisPheHisArgArgSerPheGlnSerArgPheArgGluGlnHisGlyAsnPheArg 200
QY 632 GTCTGCAAGGTTTCACCAAGGTCAAGGACGTTTTCAGAATCTCCAGAATCTCCGATTT 691
Db 201 ValLeuGlnArgPheAlaSerArgHisProIleLeuArgGlyIleAsnGluPheArgLeu 220
QY 692 GTGAGATCGAGGCCAAACCTAACACTTGTCTTCCAGCAGCTGATGCTGATAC 751
Db 221 SerIleLeuGluAlaAsnProAsnThrGluValLeuProHisCysAspAlaGluLys 240
```







Qy	1586	CGACGACCTCATCCAGTAGGCATCAACGGCTCTCTCCGAATCCCATCTGCTT-----GGC	163
Db	419	ProProGlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuLeuMetIleCys	438
Qy	1640	TTGGGTATCAACGCTGAAAAACAACACAGAAATCTTCCTTCAGGTGATAAGGACAAATGTG	1699
Db	439	PheGluValAsnAlaArgAspAsnLysLysPheThrPheIleGly---LysAspAsnIle	457
Qy	1700	ATAGACCAGATGAGAGCAAGCGAAGATTTTAGCATTTCCCTGGTCCGGTGAACAAAGTT	1759
Db	458	ValSerSerLeuAspAsnValAlaLysGluLeuAlaPheAsnTyProSerGluMetVal	477
Qy	1760	-----GAGNAGCTCATC-----	1771
Db	478	AsnGlyValPheLeuLeuGlnArgPheLeuGluArgLysLeuIleGlyArgLeuTyHis	497
Qy	1772	-----AAAAACACAGGAAGTAATCTCACTTT	1795
Db	498	LeuProHisLysAspArgLysGluSerPhePhe	508

RESULT 12

US-09-323-195A-18

; Sequence 18, Application US/09323195A  
; Patent No. 6462257  
; GENERAL INFORMATION:  
; APPLICANT: Pullman, Gerald  
; APPLICANT: Cairney, John  
; APPLICANT: Perrera, Ranjan  
; TITLE OF INVENTION: VICLIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND  
; TITLE OF INVENTION: METHODS OF USING THE SAME  
; FILE REFERENCE: IPST0009  
; CURRENT APPLICATION NUMBER: US/09/323,195A  
; CURRENT FILING DATE: 1999-06-01  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 448  
; TYPE: PRT  
; ORGANISM: Picea glauca  
US-09-323-195A-18

Alignment Scores:

Pred. NO.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

3.57e-45

559.50

50.12%

31.09%

15.71%

4

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

448

134

82

178

37

10

US-10-728-051-1 (1-2032) x US-09-323-195A-18 (1-448)

Qy	545	AGGAAGAAGACATCTCGNACAAACCCITCTACTTCCGTCAAGGCGGTTTAGCACCGC	604
Db	44	ArgGluGluGluArgGluGluAsnProTyValPheHisSerAspSerPheArgThrArg	63
Qy	605	TACGGGAACCAAAACGGTAGGATCCGGGTCTCGCAGAGGTTTGACCAAGGTCAGGCAG	664
Db	64	AlaSerSerGluAlaGlyGluIleArgAlaLeuProAsnPheGlyGluValSerGluLeu	83
Qy	665	TTTCCGAATCTCCAGATCATCCGATTGTGCAGATCGAGGCCAACCTTAACTCTTGT	724
Db	84	LeuGluGlyIleArgLysPheArgValThrCysIleGluMetLysProAsnThrValMet	103
Qy	725	CTTCCAAAGCAGCGTAGTCGTGATAACATCTTGTATCCAGCAAGGCGCAACCGGTG	784
Db	104	LeuProHisTyrlleAspAlaThrTrpIleLeuTyValThrArgGlyArgGlyTyrlle	123
Qy	785	ACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATGCATCAGA	844
Db	124	AlaTyValHisGlnAsnGluLeuValLysArgLysLeuGluGlyAspValPheGly	143
Qy	845	ATCCCATCCGGTTTCATTTCCCTACATCTTTGAACGCCCATGACCAACAGACCTCAGAGTA	904



```
QY 782 GTGACCGTAGCAATGCAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATGCACCTC 841
Db 170 LeuGlyLeuValArgGluSerGluThrGluLysIleThrLeuGluProGlyAspMetIle 189
QY 842 AGAATCCCATCCGGTTTCATTTCTACATCTTGAAACCGCCATGACACCAACCACTCAGA 901
Db 190 HisIleProAlaGlyThrProLeuTyrlleValAsnArgAspGluAsnGluLysLeuLeu 209
QY 902 GTAGCTAAATCTCCATCCCGCTTAAACACACCCCGCCAGTTTGAGGATTTCTCCCGCGC 961
Db 210 LeuAlaMetLeuHisIleProValSerThrProGlyLysPheGluGluPhePheGlyPro 229
QY 962 AGCAGCCGAGACCAATCATCTCTACTTCAGGGCTTCAGCAGGAATACGTTGGAGCGGCC 1021
Db 230 GlyGlyArgAspProGluSerValLeuSerAlaPheSerTrpAsnValLeuGlnAlaAla 249
QY 1022 TTCATCCGGAATTCATAGATACGAGGAGGTGCTGTTAGAAAGAGAAATGCAGGAGGTGAG 1081
Db 250 LeuGlnThrProLysGlyLysLeuGluArgLeuPhe 261
QY 1082 CAAGAGGAGAGCGCAGAGCGGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGTG 1141
Db 262 -----AsnGlnAsnGluGlySer 268
QY 1142 ATAGTCAAGTGTCAAGGAGCAGCTTGAAGAACTTACTAAGCAGCTAAATCGGCTCA 1201
Db 269 IlePheLysIleSerArgGluArgValArgAlaLeu-----AlaProThr 283
QY 1202 AAGAAGGCTCC-----GAAGAAGAGGGAGATATACCAACCAATCACTTGAGAGAA 1255
Db 284 LysLysSerSerTrpTrpProPheGlyGlySerLysAlaGlnPheAsnIlePheSer 303
QY 1256 GGCAGGCGCGATCTTCTAAACACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAA 1315
Db 304 LysArgProThrPheSerAsnGlyTyrglyArgLeuThrGluValGlyProAspAspGlu 323
QY 1316 AACCCCGAGCTTACGACCTGGACATGATGCTACCTGCTGAGATCAAGAAAGAGCT 1375
Db 324 LysSerTrpLeuGlnArgLeuAsnLeuMetLeuThrPheThrAsnIleThrGlnArgSer 343
QY 1376 TTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTGTCGTCGTCACAAAGGAAGT 1435
Db 344 MetSerThrIleHisTyrlleAsnSerHisAlaThrLysIleAlaLeuValMetAspGlyArg 363
QY 1436 GGAACCTTGAACCTCGTGGCTGTGAAGAAAGAGCAACAAAGAGGGGACGCGCGGAAGAA 1495
Db 364 GlyHisLeuGlnIleSerCysProHisMetSerSerArg----- 376
QY 1496 GAGGAGCAAGCAAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1555
Db 377 -----SerAspSerLysHisAspLysSerSerProSerTyrlleHisArgIleSerAla 393
QY 1556 AGGTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1600
Db 394 AspLeuLysProGlyMetValPheValProProGlyHisPro 408

RESULT 14
US-09-323-195A-17
; Sequence 17, Application US/09323195A
; Patent No. 6462257
; GENERAL INFORMATION:
; APPLICANT: Pullman, Gerald
; APPLICANT: Cairney, John
; APPLICANT: Perrera, Ranjan
; TITLE OF INVENTION: VICILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 1PST0009
; CURRENT APPLICATION NUMBER: US/09/323,195A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 523
```

```
; TYPE: PRT
; ORGANISM: Pinus taeda
US-09-323-195A-17

Alignment Scores:
Pred. No.: 1,13e-40 Length: 523
Score: 513.50 Matches: 151
Percent Similarity: 46.23% Conservative: 88
Best Local Similarity: 29.21% Mismatches: 215
Query Match: 14.42% Indels: 63
DB: 4 Gaps: 16

US-10-728-051-1 (1-2032) x US-09-323-195A-17 (1-523)
QY 533 GGTAGCATGTGTGAGGAAGAAACATCTCGAAACACCCCTTTCTACTTCCGTCGAAGCGG 592
Db 43 GlyHisGlnArgGluGluGluArgGluGluAsnProTyrlleValPheHisSerAspArg 62
QY 593 TTTAGCACCGCTACGGGAACCAACGAGTAGAGTCCGGTCTCGCAGAGGTTTGACCAA 652
Db 63 PheArgMetArgAlaSerSerAspAlaGlyGluIleArgAlaLeuProAsnPheGlyGlu 82
QY 653 AGCTCAAGCGAGTTTCCAGAATCTCCAGAATCCAGATTCACCGTATTGTGCAGATCGAGGCCAAACCT 712
Db 83 AlaSerGluLeuLeuGluGlyIleSerLysTyrlleValThrCysIleGluMetArgPro 102
QY 713 AACACTCTGTTCTTCCCAAGCAGCTGATGTGATAACATCTTGTATTCACGACGAGG 772
Db 103 AsnThrValMetLeuProHisTyrlleAspAlaThrTrpIleLeuTyrlleValThrGlyGly 122
QY 773 CAAGCCACCGTAGCCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTGAGGAGG 832
Db 123 ArgGlyTyrlleAlaTyrlleHisGlnAsnGluLeuValLysArgLysLeuGluGluGly 142
QY 833 CATGCACTCAGATCCATCCCGCTTTCATTCCTACATCTTGAAACCGCCATCAACACAG 892
Db 143 AspValPheGlyValProSerGlyHisThrPheTyrlleValAsnAsnAspAspHisAsn 162
QY 893 AACCTCAGAGTAGCTAAAATCTCCATCCCGCTTAAACACA---CCCGCCAGTTTGGAGAT 949
Db 163 SerLeuArgIleThrSerLeuLeuArgThrValSerThrMetArgGlyGluTyrllePro 182
QY 950 TTTCTCCCGGAGCAGCGCAGACCAATCATCTCTTGCAGGCTTCAGCAGGAATACG 1009
Db 183 TyrlleValAlaGlyArgAsnProGluThrValTyrlleSerAlaPheSerAspVal 202
QY 1010 TTGAGCGCGCTTCAATCGGAATTCATAGATACGAGG---GTGCTCTTA 1060
Db 203 LeuGluAlaAlaPheAsnThrAsnValIleGluAlaArgThrHisPheProValHisIle 222
QY 1061 GAAGAGAATGCAGGAGGTGAGCAAGAGAGAGAGGCGAGGCGAGTACTCGGAGT 1120
Db 223 GluArg-----GluSerTyrlleSer 228
QY 1121 AGTGAGAACATGAGGAGGTAGTCAAGTGTCAAGAGGACGTTTGAAGAACTTACT 1180
Db 229 MetAlaAsnGluGluGlnIle-----ArgGluMetLeuArg 240
QY 1181 AAGCAGCTAAAATCCGCTCAAGAAAGGCTCCGAAGAGAGGAGGAGATATCACCACCA 1240
Db 241 LysArgGlyPheSerAlaGluSerMetSerAlaSerGluHis-----ProLysPro 257
QY 1241 ATCAACTTGAGAGAAGCGCGCGATCTTCTTAACAACCTTTGGGAAGTTTATTTGAGGTG 1300
Db 258 PheAsnLeuArgAsnGlnLysProAspPheGluAsnAspAsnGlyArgPheThrArgAla 277
QY 1301 AAGCCAGACAAGAAACCCCGAGCTTCAGGACCTCGACATGATGCTCCTGCTGTAGAG 1360
Db 278 GlyProAsn---GluAsnProLeuLeuAspAlaValAspValThrAlaGlyPheGlyVal 296
QY 1361 ATCAAGAGAGGAGCTTTGATGCTCCACACTTCACTCAAGGCCATGGTTATCGTCGTC 1420
Db 297 LeuAsnProGlyThrMetThrAlaProSerHisAsnThrLysAlaThrSerIleAlaIle 316
```



```
Db 274 Phe----- 274
Qy 1115 CGGAGTAGTGAGAACATGAGAGTGTATGCTAAAGTGTCAAAGGAGCACGTTGAGAA 1174
Db 275 -----AspGlnGlnAsnGluGlySerIlePheArgIleSerArgGluGlnValArgAla 292
Qy 1175 CTTACTAAGCAGCGTAATCGGCTCAAGAGAGGCTCC-----GAAGAGAGGGAGAT 1228
Db 293 Leu-----AlaProThrLysSerSerTrpTrpPheGlyGlyGlu 307
Qy 1229 ATCACCAACCAATCAACTTGAGAGAAGCGGAGCGCATCTTTCTAACAACCTTTGGGAAG 1288
Db 308 SerLysProGlnPheAsnIlePheSerLysArgProThrIleSerAsnGlyTyrGlyArg 327
Qy 1289 TTATTGAGGTGAAGCCA--GACAAGAGAAGAACCCAGCTTCAGGACCTGAGACATGATG 1345
Db 328 LeuThrGluValGlyProAspAspGluLysSerTrpLeuGlnArgLeuAsnLeuMet 347
Qy 1346 CTCACCTGTGTAGAGATCAAGAGAGGAGCTTTGATGCTCCACACCTTCAACTCAAGGCC 1405
Db 348 LeuThrPheThrAsnIleThrGlnArgSerMetSerThrIleHisTyrAsnSerHisAla 367
Qy 1406 ATGTTTATCGTCTGTCGTCACAAAGGAACTGGAAACCTTGAACCTCGTGGCTGTAAGAAA 1465
Db 368 ThrLysIleAlaLeuValIleAspGlyArgGlyHisLeuGlnIleSerCysProHisMet 387
Qy 1466 GAGCAACAACAGAGGGGACCGCGGAGAGAGGAGGACGAGACGAAAGACGAAGAGAGGGGA 1525
Db 388 SerSerArgSerSerHisSerLysHisAspLys----- 398
Qy 1526 AGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTGAAGAGGCGATGTTCATCATG 1585
Db 399 SerSerProSerTyrHisArgIleSerSerAspLeuLysProGlyMetValPheValVal 418
Qy 1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTCCGAACCTCCATCTGCTT-----GGC 1639
Db 419 ProProGlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuLeuMetIleCys 438
Qy 1640 TTCGGTATCAACGCT 1654
Db 439 PheGluValAsnAla 443
```

Search completed: April 8, 2005, 05:34:08  
Job time : 104.925 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 8, 2005, 04:05:21 ; Search time 68.9539 Seconds

(without alignments)  
5670.817 Million cell updates/sec

Title: US-10-728-051-1

Perfect score: 3562

Sequence: 1 ataatacatatatattcatc.....cgttgtgcgtgtttctctcc 2032

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool\_p/US10728051/runat 07042005 125045 14060/app\_query.fasta\_1.4757  
-DB=PIR -QPM=fastan -SUFFIX=rdx -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bicb -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10728051@cgn 1.1 228 @runat 07042005 125045 14060 -NCPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR.79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1282	36.0	605	2	S20007
2	1281	36.0	605	1	FWSYBA
3	1267	35.6	571	2	S00566
4	1251	35.1	639	2	B24810
5	1178.5	33.1	459	2	S08505
6	1174.5	33.0	433	2	S00567
7	1156	32.5	439	1	FWSYCB
8	1148	32.2	438	2	S35757
9	1109.5	31.1	463	2	A27288
10	1108.5	31.1	463	2	S06309
11	1068	30.0	410	1	FWPMBV
12	962.5	27.0	445	2	J02264
13	959.5	26.9	445	2	S00281
14	907.5	25.5	810	2	T44430

#### ALIGNMENTS

##### RESULT 1

S20007

beta-conglycinin alpha chain precursor - soybean

C;Species: Glycine max (soybean)

C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C;Accession: S20007

R;Lelievre, J.M.; Dickinson, C.D.; Dickinson, L.A.; Nielsen, N.C.

Plant Mol. Biol. 18, 259-274, 1992

A;Title: Synthesis and assembly of soybean beta-conglycinin in vitro.

A;Reference number: S20007; MUID:92119248; PMID:1731988

A;Accession: S20007

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-605 <LEL>

A;Cross-references: UNIPROT:Q94LX2

C;Superfamily: glycinin

Alignment Scores:

Pred. No.: 8.03e-82 Length: 605

Score: 1282.00 Matches: 285

Percent Similarity: 59.37% Conservative: 108

Best Local Similarity: 43.05% Mismatches: 173

Query Match: 35.99% Indels: 96

DB: 2 Gaps: 18

US-10-728-051-1 (1-2032) x S20007 (1-605)

Qy	50	ATGAGAGGAGGGTGTTCCTCCTAGTGTCTTCTAGGATCCCTGTCTGGTTCAGTT	109
Db	2	MetArgAlaArgPhePro-----LeuLeuLeuGlyLeuValPheLeuAlaSerVal	19
Qy	110	TCGTGCAAGCATGCCAGTATCATCTTACCAGAAACAGACAGACCCCTGGCCGAG	169
Db	20	SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn	36
Qy	170	AGTGTCTCCAGATTGTCAACAGGACCGGATGACTTGAAGCAAGGATGCGAGTCT	229
Db	37	LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla	56

alpha-globulin typ  
phaseolin type alp  
vicilin precursor  
alpha-phaseolin pr  
phaseolin beta cha  
alpha-globulin typ  
alpha-globulin B p  
vicilin A precursor  
vicilin-like stora  
globulin-1S, GUB1S  
globulin Begl prec  
beta-conglycinin a  
globulin-2 precurs  
62K sucrose-bindin  
62K sucrose-bindin  
vicilin-like stora  
7S storage protein  
convicilin (clone  
vicilin-like stora  
vicilin - zania fu  
convicilin precurs  
phaseolin - kidney  
probable major pro  
vicilin, 14K compo  
globulin1 - maize  
globulin-1 - maize  
hypothetical prote  
probable seed stor  
hypothetical prote  
globulin-like prot  
beta-globulin B pr

QY	230	CGCTGCAACAAGCTCGAGTATGAT-----CCTCGTTGT	262
Db	57	ArgCysAsnLeuLeuLysValGlnLysGluGluCysGluGlyGluIleProArg---	75
QY	263	GTCATTGATCCTCGAGGACACACTGGCCACCACCAACCACTGCCCTCCAGGGGAGCGG	322
Db	76	-----ProArgProArg-----ProGlnHisProGlu	84
QY	323	ACAGTGCCCGCCAAACCCGGAGACTACGATGATGAC-----	358
Db	85	ArgGluProGlnGlnProGlyGlyLysGluGluAspGluAspGluInProArgProile	104
QY	359	-----CGCGGTCAACCCCGAAGACAG-----GAAGGAGGCCGA	391
Db	105	ProPheProArgProGlnProArgGlnGlnGluHisGluGlnArgGluGlnGln	124
QY	392	TGG-----GGACAGCTGGACCGGAGCGCTGAAAGAGAGAA	430
Db	125	TrpProArgLysGluGluLysArgGlyGlyLysGlySerGluGluGluAspGlu	144
QY	431	GACTGGAGACAAACAGAAGAT-----TGGAGGGACCAAGTCACTACGACGCCACGG	484
Db	145	AspGluGlnGlnAspGluArgGlnPheProPheProArgProProHisGlnLysGluGlu	164
QY	485	AAAATAAGCCCGAAGGAAGAGAGAACAAAGAGTGGGAAACACAGGTAGCATGTG	544
Db	165	ArgLysGlnGluGluAspGluAspGlnGlnGlnArgGluSerGluSerGluAsp	184
QY	545	AGGGAA-----GAAACATCTCGGAAACAACCCCTTTCTACTTCCCGTCAAAGCGGTTTTAGC	598
Db	185	SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu	204
QY	599	ACCCGCTACGGGAACCAAAACGGTAGAGTCCGGGTCTCTGCAGAGGTTTGACCAAGTCA	658
Db	205	ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer	224
QY	659	AGGCAGTTTCCAGATCTCCAAGATCACCGTATCTGCAGATCGAGGCCCAACCTTAACACT	718
Db	225	ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr	244
QY	719	CTTGTTCTTCCAAAGCAGCTGATGCTGATAACATCTCTGTTATTCACGACGAGGCAAGCC	778
Db	245	LeuLeuLeuProAsnHisAlaAspAlaAspTyrIleuIleValIleLeuAsnGlyThrAla	264
QY	779	ACCGTAGCGTAGCAATGGCAATAACAGAAAGACTTAACTCTGACGAGGCGGCATGCA	838
Db	265	IleLeuSerLeuValAsnAsnAspAspArgAspSerTyrArgLeuGlnSerGlyAspAla	284
QY	839	CTCAGAACTCCATCCGTTTCATTCTTACATCTTTGAACGCCCATGACACACCAAGAACCTC	898
Db	285	LeuArgValProSerGlyThrThyTyrValValAsnProAsnAsnGluAsnLeu	304
QY	899	AGAGTAGCTAAATCTCCATGCCCGTTTAACACACCCGGCCAGTTTGAGGATTTCTCCCG	958
Db	305	ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu	324
QY	959	GCGAGCAGCCGAGCAACATCATCTACTTCGAGGCGCTTCAGCAGGAATACGTTGGAGGCC	1018
Db	325	SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla	344
QY	1019	GCCTTCAATCGGAATTCATAGATACGAGGGTGCTGTAGAGAAGATGCGAGGAGGT	1078
Db	345	SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly---	363
QY	1079	GAGCAAGAGGAGGAGCGGACGCGTAGTAGTACTCGGAGTAGTGAGAACAAATGACGA	1138
Db	364	-----GlnGlnGlnGlyGluGlnArgLeu-----GlnGluSer	374
QY	1139	GTGATAGCTAAAGTGTCAAAGGACAGCTTGAAGNACTTACTAAGCAGCTTAATCCGCT	1198
Db	375	ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer	394

QY	1199	TCAAAGAAAGCGCTCCGAAAGAGAGAGGAGAGATATACCAACCCATCAACTTGAAGAGAGGC	1258
Db	395	SerArglysthrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg	411
QY	1259	GAGCCGATCTTTCTAAACAATTGGGAAGTATTATTTGAGGTGAAGCCAGACAAGAAGAAC	1318
Db	412	AspProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn	430
QY	1319	CCCAGCTTCAAGGACCTGGACATGANGCTCACCTGTGTAGAGATCAAAAGAGAGGCTTTG	1378
Db	431	ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu	450
QY	1379	ATGCTCCCACTCAACTCAAGGCCATCGTTATCGTCGTCTCAACAAGGAAGCTGGA	1438
Db	451	LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla	470
QY	1439	AACCTTGAACCTCGTGGCTGTAAAGAAAAGACAAACAAGAGGGGACGGCGGGAAGAAGAG	1498
Db	471	AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGluGluGln	489
QY	1499	GAGGACGACGACGAAAGAGGAGGGAAGTAAACAGAGAGGTGCGTAGGTACACAGCGAGG	1558
Db	490	Pro-----LeuGluValArgLysTyrArgAlaGlu	499
QY	1559	TTGAAGGAAGCGATGTTCATCATGCCACGAGCTCATCCAGTAGCCATCAACGCTTCC	1618
Db	500	LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr	519
QY	1619	TCCGAACCTCCATCTGCTTGGCTTCGCTATCAACGCTGAAAAACAACACAGAAATCTTCCCTT	1678
Db	520	SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu	539
QY	1679	GCAGGTCAATAGGACATGTGATAGACCAGATAGACGAACGACGAAAGATTAGCATTC	1738
Db	540	AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe	559
QY	1739	CCTGGGTCGGGTGAAACAAGTTGAGAAGCTCATCAAAAAACAGAAAGAAATCTCACTTTGTG	1798
Db	560	ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal	579
QY	1799	ATGTCGTCTCAATCTCAATCTCAATCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTGAG	1858
Db	580	AspAlaGlnProLysLys-----	585
QY	1859	AAAGAGATCAAGAGGAGGAAACCAGAGGGAAGGCTCCACTCTTTCAATTTTGAAG	1918
Db	586	-----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg	602
QY	1919	GTCTTT	1924
Db	603	AlaPhe	604

RESULT 2

FWSYBA

beta-conglycinin alpha chain precursor - soybean

C:Species: Glycine max (soybean)

C:Date: 31-Dec-1991 #sequence revision 31-Dec-1991 #text\_change 09-Jul-2004

C:Accession: S14681; S74124; S06714

R:Sebastiani, P.L.; Farrell, L.B.; Schuler, M.A.; Beachy, R.N.

P:Plant Mol. Biol. 15, 197-201, 1990

A:Title: Complete sequence of a cDNA of alpha subunit of soybean beta-conglycinin

A:Reference number: S14681; MUID:91355860; PMID:2103438

A: Experimental source: seed

C: Superfamily: glycinin

F: 1-22/Domain: signal sequence #status predicted &lt;SIG&gt;

F: 23-62/Domain: propeptide #status predicted &lt;PRO&gt;

F: 63-605/Product: beta-conglycinin alpha chain #status predicted &lt;WAT&gt;

F: 261,517/Binding site: carbohydrate (Asn). (covalent) #status predicted

## Alignment Scores:

Pred. No.:	9,43e-82	Length:	605
Score:	1281.00	Matches:	285
Percent Similarity:	59.21%	Conservative:	107
Best Local Similarity:	43.05%	Mismatches:	174
Query Match:	35.96%	Indels:	96
DB:	1	Gaps:	18

US-10-728-051-1 (1-2032) x FWSYBA (1-605)

Qy	50	ATGAGGGAGGTTTCTCAGTATGCTGCTAGGATCCTTGCTGGCTTCAGTT	109
Db	2	MetArgalaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal	19
Qy	110	TCTCAACGCGATCCAAAGTCATCCTTACCAGAGAAACAGAGAACCCCTCGCCCGAG	169
Db	20	SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn	36
Qy	170	AGTGCTCCAGAGTTCTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGAGTCT	229
Db	37	LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla	56
Qy	230	CGTGCACCAAGCTCGAGTATGAT-----CCTCGTTGT	262
Db	57	ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyGluIleProArg---	75
Qy	263	GTCTATGATCTCGAGGACACACTGGCACCACCAACCAACGTTCCCTCCAGGGACGG	322
Db	76	-----ProArgProArg-----ProGlnHisProGlu	84
Qy	323	ACAGTGGCGGCCCAACCGGAGACTAGCATGATGAC-----	358
Db	85	ArgGluProGlnGlnProGlyGluLysGluAspGluAspGluGlnProArgProIle	104
Qy	359	-----CGCCTCAACCCCGAAGAG-----GAAGAGGCCGA	391
Db	105	ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGluGlnGlu	124
Qy	392	TGG-----GGACCAGCTGGACGGAGGAGCGTGAAGAGAGAA	430
Db	125	TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGlu	144
Qy	431	GACTGGAGACAACCAAGAGAAGAT-----TGGAGGCGACCAAGTCATCAGACCCACGG	484
Db	145	AspGluGlnAspGluArgGlnPheProArgProPheHisGlnLysGluGlu	164
Qy	485	AAATAAGGCCGGAAG	544
Db	165	ArgAsnGluGluGluAspGluGlnGlnArgGluGluGluGluGluGluGluGluGlu	184
Qy	545	AGGAA-----GAAACATCTCGGAACAACCTTTCTACTTCCCGTCAAGCGGTTTACG	598
Db	185	SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu	204
Qy	599	ACCGCTACGGGAACCAACGGTAGGATCGGCTCGGAGAGGTTTGAACCAAGTCA	658
Db	205	ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer	224
Qy	659	AGGAGTTTCAAGATCTCCAGATCACCCTATTTCGAGATCGGAGCCCAACCTTAACACT	718
Db	225	ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr	244
Qy	719	CTTGTTCTTCCAGCGCTGATGCTGATTAACCTCTGTTATCCAGCAAGGCGCAAGCC	778
Db	245	LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleLeuAsnGlyThrAla	264

Qy	779	ACCGTGACCTAGCAATAGCAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATGCA	838
Db	265	IleLeuSerLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla	284
Qy	839	CTCAGATCCATCCCGTTTCAITTCCTACATCTTGAAACCGCATGACACCAACCTC	898
Db	285	LeuArgValProSerGlyThrThrTyrTyrValValAsnProAspAsnAsnGluLeu	304
Qy	899	AGAGTAGCTAAATATCCATGCCCTTAAACACACCCCGCCAGTTTGAAGGATTTCTCCCG	958
Db	305	ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu	324
Qy	959	GCAGAGCGAGCAGCAATCATCTTCTTCCAGGCTTACGAGGAATAGCTTGAGAGCC	1018
Db	325	SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla	344
Qy	1019	GCCTTCAATCGGGAATTCATGATGATACGAGGCTGCTTAGAGAGATGCGAGGAGT	1078
Db	345	SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly---	363
Qy	1079	GAGCAAGAGGAGAGGAGCGATGAGTACTCGGAGTAGTGAGAACAAATGAAGGA	1138
Db	364	-----GlnGlnGlnGlyGlnArgLeu-----GlnGluSer	374
Qy	1139	GTGATAGTCAAGTGTCAAAGGAGCAGTTTGAAGAACTTAAAGCAGCTTAATCCGTC	1198
Db	375	ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer	394
Qy	1199	TCAAAGAAAGCTCCGAGAGAGGAGATATCCAAACCAATCACTTGAAGAGAGGC	1258
Db	395	SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg	411
Qy	1259	GAGCCGATCTTTCTAACTTTGGAGAGTATTGAGTGAGCCAGCAACAAAGAAC	1318
Db	412	AspProIleTyrSerAsnLysLeuGlyLysPheGluIleThrProGlu---LysAsn	430
Qy	1319	CCCAGCTTCAGGACCTGGACATGCTCACCTGTGTAGAGATCAAAGAGGAGCTTTG	1378
Db	431	ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu	450
Qy	1379	ATGCTCCACACTCAACTCAAGGCCATGTTATCGTCTGTCGTCACAAAGGAGCTGGA	1438
Db	451	LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla	470
Qy	1439	AACCTTGAACTCTGGCTGTAAGAAAGAGCAACACAGAGGGGACGCGGAGAGAGAG	1498
Db	471	AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGluGln	489
Qy	1499	GAGGACGAAGCAAGAGAGGAGGAGAAAGTAAACAGAGAGGTGCTAGGTACACAGCGAGG	1558
Db	490	Pro-----LeuGluValArgLysTyrArgAlaGlu	499
Qy	1559	TTGAAGAGGCGATGTTTCATCATGTCAGAGCTCATCCAGTAGCCATCAACGCTTCC	1618
Db	500	LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValAsnAlaThr	519
Qy	1619	TCCGAACTCATCTGCTTGGCTTCCGTTATCAACGCTGAAACCAACCAAGATCTTCTT	1678
Db	520	SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu	539
Qy	1679	GCAGGTGATAGGACAATGTGATAGACCATAGAGCAAGCAAGCAAGGATTTAGCATTC	1738
Db	540	AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe	559
Qy	1739	CTTGGGTCGGGTGAACAAGTTGAGAGCTCATCAAAACCAAGAGGAATCTCATTTGTG	1798
Db	560	ProGlySerAlaAlaValGluLysLeuLysAsnGlnArgGluSerTyrPheVal	579
Qy	1799	AGTGCTCGTCTCAATCTCAATCTCCGTCGTCCTCTGAGAAAGAGTCTCTCTGAG	1858
Db	580	AspAlaGlnProLysLys-----	585



QY 1505 GAGACGAAGAAGAGGAGGAAGTAACAGACAGAGAGGTGGTAGGTACACAGCGAGGTGTAAG 1564  
 Db 450 GluAspArgLysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSer 467  
 QY 1565 GAAGGCGATGTGTTTCATCATGCCAGAGAGCTATCCAGTAGCCATCAACGCTTCCTCCGAA 1624  
 Db 468 ProGlyAspValValIleIleProAlaGlyHisProValAlaIleSerAlaSerAsn 487  
 QY 1625 CTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACACAGAATCTTCCTTGAGGT 1684  
 Db 488 LeuAsnLeuLeuGlyPheGlyIleAsnAlaLysAsnAsnGlnArgAsnPheLeuSerGly 507  
 QY 1685 GATAAGGACAATGTGATAGACCAAGTAGAGAACGAGCAAGGATTTAGCATTCCTCTGGG 1744  
 Db 508 SerAspAspAsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGly 527  
 QY 1745 TCGGGTCAACAGTTGAGAGCTCATCAAAAACCAAGAGGAATCTCACCTTTGTGAGTGTCT 1804  
 Db 528 SerSerGlnGluValAsnArgLeuIleLysAsnGlnLysGlnSerHisPheAlaSerAla 547  
 QY 1805 CGTCTTCAATCTCAATCTCCGTCGTCCTGAGAAAGAGTCTCTGAGAAAGAG 1864  
 Db 548 GluPro-----Glu 550  
 QY 1865 GATCAAGAGAGGAAACCAAGGAGGAGGTCCTCTTCAATTTTGAAGGCTTTT 1924  
 Db 551 GlnLysGluGluSerGlnArgLysArgSerProLeuSerValLeuAspSerPhe 570  
 RESULT 4  
 B24810  
 beta-conglycinin alpha chain precursor - soybean  
 C/Species: Glycine max (soybean)  
 C/Date: 07-Mar-1988 #sequence\_revision 30-Jun-1991 #text\_change 09-Jul-2004  
 C/Accession: B24810; S16337  
 R/Doyle, J.J.; Schuler, M.A.; Godette, W.D.; Zenger, V.; Beachy, R.N.; Slightom, J.L.  
 J. Biol. Chem. 261, 9228-9238, 1986  
 A/Title: The glycosylated seed storage proteins of Glycine max and Phaseolus vulgaris.  
 A/Reference number: A24810; MUID:86250867; PMID:3013879  
 A/Accession: B24810  
 A/Molecule type: DNA  
 A/Residues: 1-639 <DNP>  
 A/Cross-references: UNIPROT:P11827; GB:M13759; NID:g169928; PIDN:AA01374.1; PID:g169929  
 A/Note: the authors translated the codon GGT for residue 352 as Glu  
 R/Schuler, M.A.; Schmitt, E.S.; Beachy, R.N.  
 Nucleic Acids Res. 10, 8225-8244, 1982  
 A/Title: Closely related families of genes code for the alpha and alpha' subunits of the  
 A/Reference number: S16337; MUID:83143288; PMID:6298713  
 A/Accession: S16337  
 A/Molecule type: DNA  
 A/Residues: 361-639 <SCH>  
 C/Genetics:  
 A/Introns: 278/1; 355/3; 382/3; 481/3; 575/1  
 C/Superfamily: glycinin  
 C/Keywords: seed; storage protein  
 Alignment Scores:  
 Pred. No.: 1,2&-79 Length: 639  
 Score: 1251.00 Matches: 283  
 Percent Similarity: 58.75% Conservative: 120  
 Best Local Similarity: 41.25% Mismatches: 173  
 Query Match: 35.12% Indels: 110  
 DB: 2 Gaps: 21  
 US-10-728-051-1 (1-2032) x B24810 (1-639)  
 QY 50 ATGAGAGGAGGGGTTTCTCCACTGATCTGTCTAGGGATCTTGTCTCGGCTTCAGTT 109  
 Db 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValPheValPheLeuAlaSerVal 19  
 QY 110 TCTGCAACGGATGCCAAGTCAATCATCTTACAGAGAAGAAAACAGAGAACCCCTCGGCCAC 169  
 Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GlnAsnProSerHisAsn 36

Db 397 GluGlyGlnGlnGlnGlnGluGluArgLeuGln----- 407  
QY 1127 AACAAATGAGAGGTAGTCAAGGTCAAGGAGCAGCGTGAAGAACTTACTAAGCAC 1186  
Db 408 -----GluSerValIleValGluIleSerLysGlnIleArgGluLeuSerLysHis 425  
QY 1187 GCTAAATCCGCTCTCAAGAAAGGCTCCGAAGAGAGGAGATATACCAACCCCAATCAAC 1246  
Db 426 AlaLysSerSerArgLysThrIleSerSerGlu-----AspLysProPheAsn 442  
QY 1247 TTGAGAGAGGCGAGCGGATCTTTCTCAACACTTTGGGAAGTTATTTGAGGTGAAGCCA 1306  
Db 443 LeuGlySerArgAspProIleTyrSerAsnLysLeuGlyLysLeuPheGluIle----- 460  
QY 1307 GACAAAGAACCCAGCTTCAGGACCTGGACATGATCTACCTGTGTAGAGATCAAA 1366  
Db 461 ThrGlnArgAsnProGlnLeuArgAspLysPheAsnValPheLeuSerValValAspMetAsn 480  
QY 1367 GAAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCGTCAAC 1426  
Db 481 GluGlyAlaLeuPheLeuProHisPheAsnSerLysAlaIleValValLeuValIleAsn 500  
QY 1427 AAGGAACTGGAACTTGAACCTGCTGGCTGTGAAGAAAGAGCAACACAGAGGGACGG 1486  
Db 501 GluGlyGluAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnArgGlnGln 519  
QY 1487 CGGGAAGAGAGGAGGACGAAGACGAAGAGAGGAGGAAGTAAACAGAGAGGTGCGTAGG 1546  
Db 520 GlnGluGluGlnPro-----LeuGluValArgLys 529  
QY 1547 TACACAGCGAGGTGAAGGAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCC 1606  
Db 530 TyrArgAlaGluLeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValMet 549  
QY 1607 ATCAACGCTTCCCGAATCCATCTGCTTGGCTTCGGTATCAACGCTGAACACCAACCAC 1666  
Db 550 ValAsnAlaThrSerAspLeuAsnPhePheAlaPheGlyIleAsnAlaGluAsnAsnGln 569  
QY 1667 AGAATCTTCTCGAGGTGATAGGACAATGTGATAGACACAGATAGAGGAGCAAGGCAAG 1726  
Db 570 ArgAsnPheLeuAlaGlySerLysAspAsnValIleSerGlnIleProSerGlnValGln 589  
QY 1727 GATTAGCATTCCTGGTGGTGGTGAACAAGTTGAGAGCTCATCAAAACCAAGAGGAA 1786  
Db 590 GluLeuAlaPheProArgSerAlaLysAspIleGluAsnLeuIleLysSerGlnSerGlu 609  
QY 1787 TCTCACTTTGAGTGTGCTGCTCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAA 1846  
Db 610 SerTyrPheValAspAlaGlnProGln----- 618  
QY 1847 GAGTCTCTCGAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1906  
Db 619 -----GlnLysGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 632  
QY 1907 TCAATTTTGAAGGCTTTT 1924  
Db 633 SerIleLeuArgAlaPhe 638  
RESULT 5  
S08505  
vicilin - garden pea  
C:Species: Pisum sativum (garden pea)  
C:Date: 02-Dec-1993 #sequence\_revision 01-Dec-1995 #text\_change 09-Jul-2004  
C:Accession: S08505  
P:Higgins, T.J.V.; Newbigin, E.J.; Spencer, D.; Llewellyn, D.J.; Craig, S.  
Plant Mol. Biol. 11, 683-695, 1988  
A:Title: The sequence of a pea vicilin gene and its expression in transgenic tobacco pla  
A:Reference number: S08505  
A:Accession: S08505  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-459 <HIG>

A;Cross-references: UNIPROT:P13918; EMBL:X14076; NID:G20917; PIDN:CAA32239.1; PID:G313671  
C:Genetics:  
A;Introns: 102/1; 160/3; 187/3; 295/3; 400/1  
C;Superfamily: glycinin  
Alignment Scores: 1.33e-74 Length: 459  
Pred. No.: 1178.50 Matches: 223  
Score: 71.33% Conservative: 93  
Best Local Similarity: 50.34% Mismatches: 114  
Query Match: 33.09% Indels: 13  
DB: 2 Gaps: 4  
US-10-728-051-1 (1-2032) x S08505 (1-459)  
QY 542 GTGAGGGAAGAACATCTCGGAACACCCCTTTCTACTTCCGTCAGGCGGTTTAGCAC 601  
Db 26 ValSerSerArgSerAspProGlnAsnProPheIlePheLysSerAsnLysPheGlnThr 45  
QY 602 CGCTACGGGAACCAAAACGGTAGATCCGGTCTCGCAGAGGTTTGACCAAGGTCAAGG 661  
Db 46 LeuPheGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspGlnArgSerLys 65  
QY 662 CAGTTTCAGAACTCCAGAACTCACCGTATGTGCAGATCGAGGCCAAACCTAACACTCT 721  
Db 66 IlePheGluAsnLeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrIle 85  
QY 722 GTTCTTCCCAACGACGCTGATCTGATTAACATCTCTTGTATCCAGCAAGGCGCACCC 781  
Db 86 PheLeuProGlnHisThrAspAlaAspTyrIleLeuValValLeuSerGlyLysAlaIle 105  
QY 782 GTGACCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGACAGAGGCGCATGCAC 841  
Db 106 LeuThrValLeuLysProAspAspArgAsnSerPheAsnLeuGluArgGlyAspThrIle 125  
QY 842 AGAATCCCATCGGTTTCAATTTCTTACATCTTGAACCGCATGACACACAGAACCTCAGA 901  
Db 126 LysLeuProAlaGlyThrIleAlaTyrLeuValAsnArgAspAsnGluGluLeuArg 145  
QY 902 GTAGCTTAAATCTCCATGCCCTTAAACACACCCGCCAGTTTGAGAGTTTTCGCGGG 961  
Db 146 ValLeuAspLeuAlaIleProValAsnArgProGlyGlnLeuGlnSerPheLeuLeuSer 165  
QY 962 AGCAGCGGAGACCAATCATCTTACTTGAGGCGCTTACAGAGGAATACGTTGGAGCGGCC 1021  
Db 166 GlyAsnGlnAsnGlnGlnAsnTyrLeuSerGlyPheSerLysAsnIleLeuGluAlaSer 185  
QY 1022 TTCAATGCGGAATTCATGAGATACGAGGCGTGTGTAGAGAGAATCGAGAGGTGAG 1081  
Db 186 PheAsnThrAspTyrGluGluIleGluLysValLeuLeuGluGluHisGluLysGluThr 205  
QY 1082 CAAGAGGAGAGGCGCAGAGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGTG 1141  
Db 206 GlnHisArgArgSerLeuLys-----AspLysArgGlnGlnSerGlnGluAsnVal 223  
QY 1142 ATAGTCAAAAGTGTCAAAGGAGCAGCGTTGAAGAACTTACTTAAGCAGCTAAATCCGTC 1201  
Db 224 IleValLysLeuSerArgGlyGlnIleGluLeuLeuSerLysAsnAlaLysSerThrSer 243  
QY 1202 AAGAAAGGCTCCGGAAGAGGAGATATACCAACCCCAATCAACTTGAAGAGGCGAG 1261  
Db 244 LysLysSerValSerSerGluSerGlu-----ProPheAsnLeuArgSerArgGly 260  
QY 1262 CCGATCTTTTAAACAACTTTGGGAAGTTATTGAGGTGAAGCAGCAGCAAGAGAACCC 1321  
Db 261 ProIleTyrSerAsnGluPheGlyLysPhePheGluIleThrProGlu---LysAsnPro 279  
QY 1322 CAGCTTCAGACCTCGACATGATGCTCACTGTGTAGAGATCAAAAGAGGAGCTTTGATG 1381  
Db 280 GlnLeuGlnAspLeuAspIlePheValAsnSerValGluIleLysGluGlySerLeuLeu 299  
QY 1382 CTCCACACTTCAACTCAAAGGCGCATGTTATCGTCGTCTCAACAAAGAACTGGAAC 1441



```
Db 300 LeuProHisTyrAsnSerArgAlaIleValIleValThrValAsnGluGlyLysGlyAsp 319
Qy 1442 CTTGAACCTCGTGTGTAAGAAAGAGCAACAACAGAGGGAGCGCGGGAAGAGGAG 1501
Db 320 PheGluLeuValGlyGlnArgAsnGluAsnGlnGlnGluArgLysGluAspAspGlu 339
Qy 1502 GACGACAGCAAGAGAGGAGGAGGAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTG 1561
Db 340 GluGluGlnGlnGlyGluGluGluIleAsnLysGlnValGlnAsnTyrLysAlaLysLeu 359
Qy 1562 AAGGAGCGCATGTGTTTCATCCAGCCAGAGCTCATCCAGTACGATCAACAGCTTCCTCC 1621
Db 360 SerSerGlyAspValPheValIleProAlaGlyHisProValAlaValLysAlaSerSer 379
Qy 1622 GAATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1681
Db 380 AsnLeuAspLeuLeuGlyPheGlyIleAsnAlaGluAsnGlnArgAsnPheLeuAla 399
Qy 1682 GGTGATAGACATGTGATAGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741
Db 400 GlyAspGluAspAsnValIleSerGlnIleGlnArgProValLysGluLeuAlaPhePro 419
Qy 1742 GGGTCGGGTGAACAAAGTTGAGAAGCTCATCAAAACCCAGAAAGGAATCTCACTTTGTGAGT 1801
Db 420 GlySerAlaGlnGluValAspArgIleLeuGluAsnGlnLysGlnSerHisPheAlaAsp 439
Qy 1802 GCTGCTCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 1861
Db 440 AlaGlnProGlnGlnArg-----GluArgGlySerArgGluThr 452
Qy 1862 GAGGATCAA 1870
Db 453 ArgAspArg 455

RESULT 6
S00567
vicilin precursor (clone pDUB9) - garden pea (fragment)
C;Species: Pisum sativum (garden pea)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: S00567
R;Watson, M.D.; Lambert, N.; Delaunev, A.; Yarwood, J.N.; Crox, R.R.D.; Gatehouse, J.A.;
Biochem. J. 251, 857-864, 1988
A;Title: Isolation and expression of a pea vicilin cDNA in the yeast Saccharomyces cerevisiae
A;Reference number: S00567; MUID:88326226; PMID:3046604
A;Accession: S00567
A;Molecule type: mRNA
A;Residues: 1-433 <MAT>
A;Cross-references: UNIPROT:P13918; EMBL:Y00722; NID:q20915; PIDN:CAA68708.1; PID:g75824
C;Superfamily: glycinin
F;1-2/Domain: signal sequence (fragment) #status predicted <SIG>
F;3-433/Product: vicilin #status predicted <MAT>

Alignment Scores:
Pred. No.: 2 5e-74 Length: 433
Score: 1174.50 Matches: 221
Percent Similarity: 72.18% Conservative: 93
Best Local Similarity: 50.80% Mismatches: 108
Query Match: 32.97% Indels: 13
DB: 2 Gaps: 4

US-10-728-051-1 (1-2032) x S00567 (1-433)

Qy 566 AACCTTTCTACTCCGCTCAAGCGGTTTAGCACCGCTACCGGAACCAAAACGGTAGG 625
Db 8 AsnProPheIlePheLysSerAsnLysPheGlnThrLeuPheGluAsnGluAsnGlyHis 27
Qy 626 ATCCGGCTCTCGAGAGGTTTGACCAAGGTCAAGGAGGTTTCAGATCTCCAGATCAC 685
Db 28 IleArgLeuLeuGlnLysPheAspGlnArgSerLysIlePheGluAsnLeuGlnAsnTyr 47
Qy 686 CGTATTGTGAGATCGAGGCGCAACCTTAACACTCTTGTCTTCCCAAGCAGCGTGATGCT 745
Db 48 ArgLeuLeuGluTyrLysSerLysProHisThrIlePheLeuProGlnHisThrAspAla 67
```

```
Qy 746 GATAACATCTTTGTTATTCAGCAAGGCAAGCCACCGTGACCGTAGCAATGGCAATAAC 805
Db 68 AspTyrIleLeuValLeuSerGlyLysAlaIleLeuThrValLeuLysProAspAsp 87
Qy 806 AGAAGAGCTTTAATCTTGACGAGGCGCATGCACTCAGATCCCATCCGCTTCATTTCC 865
Db 88 ArgAsnSerPheAsnLeuGluArgGlyAspThrIleLysLeuProAlaGlyThrIleAla 107
Qy 866 TACATCTTGAACCGCCATCAACACAGAACCTCAGAGTAGCTAAAATCTCCATGCCCGTT 925
Db 108 TyrLeuValAsnArgAspAsnGluLeuArgValLeuAspLeuAlaIleProVal 127
Qy 926 AACACACCGCGCGCTTGAGGATTTCTCCCGGCGAGCAGCGAGCAATCATCTCTAC 985
Db 128 AsnArgProGlyGlnLeuGlnSerPheLeuLeuSerGlyAsnGlnAsnGlnAsnTyr 147
Qy 986 TTCGAGGCTTTCAGCAGGATACGTTGGAGCGCTTCAATCGGGAATTCATGAGATA 1045
Db 148 LeuSerGlyPheSerLysAsnIleLeuGluAlaSerPheAsnThrAspTyrGluGluIle 167
Qy 1046 CGAGGGTGTGTAGAGAGAAATGCAGAGGTGAGCAAGAGAGAGAGCGGAGCGGA 1105
Db 168 GluLysValLeuLeuGluHisGluLysGluThrGlnHisArgArgSerLeuLys--- 186
Qy 1106 TGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGTAGTATGTAAGATGTCAAAGGAGCAC 1165
Db 187 ---AspLysArgGlnGlnSerGlnGluAsnValIleValLysLeuSerArgGlyGln 205
Qy 1166 GTTGAGAACTTACTAAGCAGCGTAATCCGTCTCAAGAAAGGCTCCGAAGAAGAGGGA 1225
Db 206 IleGluLeuLeuSerLysAsnAlaLysSerThrSerLysLysSerValSerSerGluSer 225
Qy 1226 GATATCACCAACCAATCACTTGAGAGAGCGGCGCCATCTTTTCAACACTTTGGG 1285
Db 226 Glu-----ProPheAsnLeuArgSerArgGlyProIleTyrSerAsnGluPheGly 242
Qy 1286 AAGTTATTGAGTGAAAGCAGACAGAAAGAACCCAGCTTCAGGACCTTGAGCATATG 1345
Db 243 LysPhePheGluIleThrProGlu---LysAsnProGlnLeuGlnAspLeuAspIlePhe 261
Qy 1346 CTCACCTGTGTAGATCAAGAGAGGCTTTCATGCTCCACACTTCACTCAAAAGGCC 1405
Db 262 ValAsnSerValGluLysGluLysLeuLeuLeuProHisTyrAsnSerArgAla 281
Qy 1406 ATGTTATCTGCTCGTCAACAAAGAACTTGAACCTCGTGGCTGTGAAGAAA 1465
Db 282 IleValIleValThrValAsnGluGlyLysGlyAspPheGluLeuValGlyGlnArgAsn 301
Qy 1466 GAGCAACAACAGAGGGAGCGCGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1525
Db 302 GluAsnGlnGlnGluGlnArgLysGluAspAspGluGluGluGlnGlyGluGluGlu 321
Qy 1526 AGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAGAAAGCGGCTGTTTCATCATG 1585
Db 322 IleAsnLysGlnValGlnAsnTyrLysAlaLysLeuSerSerGlyAspValPheValIle 341
Qy 1586 CCAGCAGCTCATCCAGTAGCCATCAACCGCTTCCTCCGAACTCCATCTGCTTGGCTTCG 1645
Db 342 ProAlaGlyHisProValAlaLeuLysAlaSerSerAsnLeuAspLeuLeuGlyPheGly 361
Qy 1646 ATCAACGCTGAAAAACAACACAGAAATCTTCTTGCAGGTGATTAAGGCAATGTGATAGAC 1705
Db 362 IleAsnAlaGluAsnAsnGlnArgAsnPheLeuAlaGlyAspGluAspAsnValIleSer 381
Qy 1706 CAGATAGAGAGCAAGCGAGGATTTAGCATTTCCCTGGGTCCGGTGAACAAGTTGAGAAG 1765
Db 382 GlnValGlnArgProValLysGluLeuAlaPheProGlySerAlaGlnGluValAspArg 401
Qy 1766 CTCATCAAAAACCAAGAGGAATCTCACTTTGTGAGTCTGCTCTCAATCTCAATCTCAA 1825
Db 402 IleLeuGluAsnGlnLysGlnSerHisPheAlaAspAlaGlnProGlnGlnArg----- 419
```

```
QY 1826 TCTCCGTCGTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGATCAA 1870
Db 420 -----GluArgGlySerArgGluThrArgAspArg 429
RESULT 7
FWSYCB
beta-conglycinin beta chain - soybean
C;Species: Glycine max (soybean)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C;Accession: JQ0969
R;Harada, J.J.; Barker, S.J.; Goldberg, R.B.
Plant Cell 1, 415-425, 1989
A;Title: Soybean beta-conglycinin genes are clustered in several DNA regions and are reg
A;Reference number: JQ0969; MUID:93005638; PMID:2562562
A;Accession: JQ0969
A;Molecule type: DNA
A;Residues: 1-439 <HAR>
A;Cross-references: UNIPROT:P25974; GB:S44893; NID:G256426; PIDN:AAB23463.1; PID:G256427
C;Comment: This protein accumulates during seed development and is hydrolyzed after germ
C;Genetics:
A;Gene: CG-4
A;Introns: 101/1; 159/3; 186/3; 281/3; 375/1
C;Superfamily: Glycinin
C;Keywords: seed; storage protein
Alignment Scores:
Pred. No.: 4,95e-73 Length: 439
Score: 1156.00 Matches: 237
Percent Similarity: 69.98% Conservative: 87
Best Local Similarity: 51.19% Mismatches: 85
Query Match: 32.45% Indels: 54
DB: 1 Gaps: 10
US-10-728-051-1 (1-2032) x FWSYCB (1-439)
QY 542 GTGAGGAAGAATCATCTCGAACAACCTTTCTACTTC---CCGTCAAGGCGGTTTACG 598
Db 26 ValArgGluAsp-----GluAsnAsnProPheTyrPheArgSerSerAsnSerPheGln 43
QY 599 ACCGCTACGGGAACCAACGTTAGGATCCGGTCTCTGAGAGGTTTGACCAAGAGTCA 658
Db 44 ThrLeuPheGluAsnGlnAsnValArgLeuGluLeuGlnArgPheAsnLysArgSer 63
QY 659 AGGCAGTTTCAGAAATCTCCAGAAATCAACGTTATGTGAGATCGAGGCCAAACCTAACAT 718
Db 64 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 83
QY 719 CTGTGTTCTCCAGGACGCTGATGCTGATAACATCTCTTCTTATCCAGCAAGGCGAGCC 778
Db 84 IleLeuLeuProHisHisAlaAspAlaAspPheLeuLeuPheValLeuSerGlyArgAla 103
QY 779 ACCGTGACCTAGCAATGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCA 838
Db 104 IleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuHisProGlyLysAla 123
QY 839 CTCAGATCCCATCCGGTTTCATTCTCATATCTTGACCCGCTATGACCAACCAAGACCTC 898
Db 124 GlnArgIleProAlaGlyThrThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 143
QY 899 AGAGTACTAAATCTCCATCCGCTTAAACACACCCGCGCAGTTTGAGGATTTCTCCCG 958
Db 144 LysIleLeuLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPheLeu 163
QY 959 GCAGAGCGCGAGACCAATCATCTCTGAGGCTTCAGAGGATACAGTTGGAGGCC 1018
Db 164 SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 183
QY 1019 GCCTTCATCGGGAATTCATATGATACGAGGCTGCTGTGTAGAAAGAAATGCGAGGAGT 1078
Db 184 SerPheHisSerGluPheGluGluIleAsnArgValLeuPhe-----Gly 198
QY 1079 GAGCAAGAGGAGGAGGCGGATGCGAGTACTCGAGTAGTGTGAGACAATGAGGA 1138
|||||
```

```
Db 199 GluGluGluGluGlnArgGlnGln-----GluGly 208
QY 1139 GTGATAGTCAAAAGTGTCAAAGGAGCAGCGTTGAAGAACTTACTAAGCAGCGTAAATCCGTC 1198
Db 209 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 228
QY 1199 TCAAAAGAAA-----GGCTCCGAAGAAGAGGAGATATCATCAACCCCAATCAACTTGAGA 1252
Db 229 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 243
QY 1253 GAAGGCGGAGCCGATCTTCTTAACAACCTTGGGAAGTTATTGAGGTGAAGCCAGACAAG 1312
Db 244 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPhePheGluIleThrProGlu--- 262
QY 1313 AAGAAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAGA 1372
Db 263 LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 282
QY 1373 GCTTTGATGCTCCACACTTCAACTCAAGGCGCATGTTATCTGCTGTCGTCACAAAGGA 1432
Db 283 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 302
QY 1433 ACTGGAAACCTTGAACCTGCTGCTGTAAAGAAAGACACACAGAGGGGACGGGGGAA 1492
Db 303 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnGlnLysGlnGlu 321
QY 1493 GAAGAGGAGGACCAAGAAGAGAGGAGGAGGAGTAAACAGAGAGGTGCTAGGTACACA 1552
Db 322 GluGluPro-----LeuGluValGlnArgTyrArg 331
QY 1553 GCGAGGTTGAAGAAAGGCGATGTGTCATCATCGCAGAGCTCATCCAGTAGCCATCAAC 1612
Db 332 AlaGluLeuSerGluAspAspValPheValIleProAlaAlaTyrProPheValValAsn 351
QY 1613 GCTTCTCCGAATCCATCTGCTTCTGCTTCAACGCTGATCAACGCTGAAACCAACCAAGAATC 1672
Db 352 AlaThrSerAsnLeuAsnPheLeuAlaPheGlyIleAsnAlaGluAsnAsnGlnArgAsn 371
QY 1673 TTCTCTGCAAGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAAGCAAGCGAAGATT 1732
Db 372 PheLeuAlaGlyGluLysAspAsnValValArgGlnIleGluArgGlnValGlnGluLeu 391
QY 1733 GCATTCCTCGGTGGGTGCAACAAGTTGAGAAGCTCATCAAAACCAAGAGGAATCTCAC 1792
Db 392 AlaPheProGlySerAlaGlnAspValGluArgLeuLeuLysLysGlnArgGluSerTyr 411
QY 1793 TTTGTGAGTCTGCTCTCAATCTCAATCTCAATCTCCGCTCTCTCTGAGAAAGAGTCT 1852
Db 412 PheValAspAlaGlnProGln----- 418
QY 1853 CTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1912
Db 419 -----GlnLysGluGlySerLysGlyArgLysGlyPropheProSerIle 434
QY 1913 TTGAAGGCT 1921
Db 435 LeuGlyAla 437
RESULT 8
S35757
vicilin, 47K - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S35757
R;Bown, D.
submitted to the EMBL Data Library, June 1992
A;Reference number: S35747
A;Accession: S35757
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-438 <BOW>
A;Cross-references: UNIPROT:Q43626; EMBL:X67429; NID:G297169; PIDN:CAA47814.1; PID:G29717
C;Superfamily: glycinin
```

110 LeuProAsnAspArgAsnSerPheSerLeuGluArgGlyAspThrIleLysLeuProAla 129

[illegible]

QY 1034 TTCAATGAGATACGAGGGTCTGTTAGAGAGAAATCAGAGAGTGTAGCAGAGCAGAGA 1093  
Db 190 TyrlsgluilleGluLysValLeuLeuGluGluHisGlyLysGluLysTyrHisArgArg 209  
QY 1094 GGGCAGAGCGATGAGTACTCGAGTAGTGTAGAACAAATGAGGAGTGTAGTCAAGTG 1153  
Db 210 GlyLeuLysAspArgArgGlnArgGlyGlnGluLeuAsn-----ValIleValLysIle 227  
QY 1154 TCAAGAGGACGCGTTGAAGAACTTACTAAGCACGCTAAATCCGCTCAAGAAAGGCTCC 1213  
Db 228 SerArgLysGlnIleGluLeuAsnLysAsnAlaLysSerSerLysLysSerThr 247  
QY 1214 GAGAGAGGAGATATCACCACCAATCAACTTGTAGAGAGCGCGAGTCTTTCT 1273  
Db 248 SerSerGluSerGlu-----ProPheAsnLeuArgSerArgGluProIleTyrSer 264  
QY 1274 AACAACTTTCGGAGGTATTGTAGGTCAAGCCAGAGAACACCCCGCTTCAGGAC 1333  
Db 265 AsnLysPheGlyLysPhePheGluIleThrPro----LysArgAsnProGlnLeuGlnAsp 283  
QY 1334 CTGACATGATGCTCACCTGTGTAGATCAAGAGGAGCTTTGATGCTCCACACTTC 1393  
Db 284 LeuAsnIlePheValAsnTyrValGluIleAsnGluGlySerLeuLeuLeuProHisTyr 303  
QY 1394 AACTCAAGCGCATGTTTGTGCTGCTGCTCAACAAAGAACTTGAACCTTGAACCTG 1453  
Db 304 AsnSerArgAlaIleValIleValThrValAsnGluGlyLysGlyAspPheGluLeuVal 323  
QY 1454 GCTGTAGAAAGAGCAACACAGAGGGGCGCGGAGAGAGAGGAGGAGAGAGCAA 1513  
Db 324 GlyLysArgAsnGluAsnGlnGlnGlyLeuArgGluGluTyrAspGluGluLysGluGln 343  
QY 1514 GAAGAGAGGAGGAGTACAGAGAGTGTGTAGGTATACAGAGGTTGAAGGAGGAGAT 1573  
Db 344 GlyGluGluGluIleArgLysGlnValGlnAsnTyrLysAlaLysLeuSerProGlyAsp 363  
QY 1574 GTGTTATCATGCGAGCAGCTCATCCAGTACCCATCAACGCTTCTCCGAACTCCATCTG 1633  
Db 364 ValLeuValIleProAlaGlyTyrProValAlaIleLysAlaSerSerAsnLeuAsnLeu 383  
QY 1634 CTGGCTTCGGTATCAGCGTGAACAAACACAGATCTCTCTGAGGTGTAGGAC 1693  
Db 384 ValGlyPheGlyIleAsnAlaGluAsnAsnGlnArgTyrPheLeuAlaGlyGluAsp 403  
QY 1694 AATGTGTATAGACAGATAGAGAGCAAGCAAGGATTTAGCATTCCTCGGTGGGTGAA 1753  
Db 404 AsnValIleSerGlnIleHisLysProValLysGluLeuAlaPheProGlySerAlaGln 423  
QY 1754 CAAGTTGAGAGGTCTATCAAAACCAGAGAAATCTCAGTGTGTAGTGTCTCTCA 1813  
Db 424 GluValAspThrLeuLeuGluAsnGlnLysGlnSerHisPheAlaAsnAlaGlnProArg 443  
QY 1814 TCTCAATCTCAATCTCGTGTCTCTGAGAGAGAGTCTCTGAGAAAGAGATCAAGAG 1873  
Db 444 -----GluArgGlu 446  
QY 1874 GAGGAAACCAAGAGGAGGAGGTCCTCTTCAATTTTGAAGGCTTT 1924  
Db 447 ArgGlySerGlnGluIleLysAspHisLeuTyrSerIleLeuGlySerPhe 463

RESULT 11  
FWPMVB  
vicilin B precursor - garden pea (fragment)  
C:Species: Pisum sativum (garden pea)  
C:Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 09-Jul-2004  
R:Lycett, G.W.; Delauney, A.J.; Gatehouse, J.A.; Gilroy, J.; Croy, R.R.D.; Boulter, D.  
Nucleic Acids Res. 11, 2367-2380, 1983  
A:Title: The vicilin gene family of pea (Pisum sativum L.): a complete cDNA coding sequence  
A:Reference number: A93462; MUID:83220791; PMID:6687941  
A:Accession: A03344  
A:Molecule type: mRNA

A;Residues: 1-410 <LYC>  
A;Cross-references: UNIPROT:P02854  
A;Experimental source: cv. Feltham First, clones pDUB7 and pDUB4  
A;Note: parts of this sequence, including the amino end of the mature protein, were determined from a cDNA library.  
C;Comment: The gene that codes for this protein is part of a multigene family coding for vicilin B precursor.  
C;Superfamily: glycinin  
C;Keywords: glycoprotein  
F;1-15/Domain: signal sequence #status predicted -SIG>  
F;16-410/Product: vicilin type B (fragment) #status predicted <MAT>  
F;321-322/Cleavage site: Asn-Asp (unidentified proteinase) #status experimental  
F;359/Binding site: carbohydrate (Asn) (covalent) #status experimental

Alignment Scores:  
Pred. No.: 7,05e-67 Length: 410  
Score: 1068.00 Matches: 209  
Percent Similarity: 71.64% Conservative: 84  
Best Local Similarity: 51.10% Mismatches: 104  
Query Match: 29.98% Indels: 12  
DB: 1 Gaps: 5

US-10-728-051-1 (1-2032) x FWPMB (1-410)  
QY 542 GTGAGGAGAGAAATCTCGGAACAACCTTTCTACTTCCGTCAGGCGGTTTAGCACC 601  
Db 13 ValSerSerArgSerAspGlnGluAsnProPheIlePheLysSerAsnArgPheGlnThr 32  
QY 602 CGCTACGGGAACCAAAACGCTAGGATCCGGGTCTCTGAGAGGTTTGACCAAGGTCAAGG 661  
Db 33 LeuTyrGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspLysArgSerLys 52  
QY 662 CAGTTTCAGAACTCTCAGAAATCAGTATGTCAGATCGAGGCCAAACCTTAACACTCTT 721  
Db 53 IlePheGluAsnLeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrLeu 72  
QY 722 GTTCTTCCCAAGCAGCTGTGTGTATACATCTCTGTTATCCAGAGGCAAGCCACC 781  
Db 73 PheLeuProGlnTyrThrAspAlaAspPheIleLeuValLeuSerGlyLysAlaThr 92  
QY 782 GTGACGCTAGCAATGCAATAACAGAGAGCTTTAATCTTCAGAGGCGCATGCACTC 841  
Db 93 LeuThrValLeuLysSerAsnAspArgAsnSerPheAsnLeuGluArgGlyAspAlaIle 112  
QY 842 AGAATCCCATCCGGTTTCATTTCTTACATCTTGAACCGCCATGACAAACAGACCTCAGA 901  
Db 113 LysLeuProAlaGlySerIleAlaTyrPheAlaAsnArgAspAsnGluGluProArg 132  
QY 902 GTAGCTAAATCTCCATGCGCTTAACACACCGCGCAGTTTTCAGGATTTCTCCCGCG 961  
Db 133 ValLeuAspLeuAlaIleProValAsnLysProGlyGlnLeuGlnSerPheLeuLeuSer 152  
QY 962 AGCAGCGAGACCAATCATCTACTTTCAGGCGCTTCAGAGGAAATAGCTTTCGAGCGCGCC 1021  
Db 153 GlyThrGlnAsnGlnLysSerSerLeuSerGlyPheSerLysAsnIleLeuGluAlaAla 172  
QY 1022 TTCATTCGGAAATTCATAGATACGAGGCGTCTGTGTAGAGAGAAATGCGAGGTTGAG 1081  
Db 173 PheAsnThrAsnTyrGluGluIleGluLysValLeuLeuGluGln-----Glu 189  
QY 1082 CRAGAGAGAGAGCGGAGCGCATGGAGTACTCGGAGTAGTAGTAC---AACATGAGGA 1138  
Db 190 GlnGluProGlnHisArgSerLeuLysAspArgGlnGluIleAsnGluGluAsn 209  
QY 1139 GTGATAGTCAAGGTCAAGGAGGACGCTTGAAGAACTTACTAAGCACGCTTAATCCGTC 1198  
Db 210 ValIleValLysValSerArgAspGlnIleGluLeuSerLysAsnAlaLysSerSer 229  
QY 1199 TCAAGAAAGCGCTCCGAGAGAGGAGATATACCAACCCCAATCAACTTTCAGAGAGGC 1258  
Db 230 SerLysLysSerValSerSerGlu-----SerGlyProPheAsnLeuArgSerArg 246  
QY 1259 GAGCCCGCATCTTCTAACAACCTTTCGAGAGTATTTGAGGTGAAGCCAGACAGAAAC 1318  
Db 247 AsnProIleTyrSerAsnLysPheGlyLysPhePheGluIleThrProGlu---LysAsn 265





Pred. No.:	2.82e-59	Length:	445
Score:	959.50	Matches:	197
Percent Similarity:	54.03%	Conservative:	118
Best Local Similarity:	33.79%	Mismatches:	101
Query Match:	26.94%	Indels:	167
DB:	2	Gaps:	8
US-10-728-051-1 (1-2032) x S00281 (1-445)			
Qy	68	CCACTGATGCTGTGTCTAGGATCCTTGCTCGCTTCAGTTCTGTGCAACGCATGCCAAG	127
Db	8	ProLeuTrpLeuLeuGlyValValLeuLeuAlaSerValSerAlaSerPheAlaHis	27
Qy	128	TCATCCTTACCCAGAGAAAAAGAGAACCCCTCGCCAGAGGTGCTCCAGAGTTGT	187
Db	28	Ser-----	28
Qy	188	CAACAGAACCGGATGACTTGAACGCAAAAGCATGCGAGTCTCGCTGCCAACAGCTCGAG	247
Db	28	-----	28
Qy	248	TATGATCCTCGTTGTGTCTATGATCCTCGAGGAGACACTGGCACCACCAACGATGCC	307
Db	29	-----GlyHisSerGly-----	32
Qy	308	CCTCCAGGGAGCGGACACGCTGGCCGCCAACCCCGAGACTACGATGATGACCGCGTCAA	367
Db	33	-----GlyGluAlaGluAspGluSerGluGlu	41
Qy	368	CCCGAAGAGAGAAAGGAGGCCGATGGGACACAGCTGGACGGAGGAGCGTGAAGAGAA	427
Db	42	SerArg-----	43
Qy	428	GAAGACTGGAGACAAACAGAGAAGATTGGAGCGGACCAAGTCATCAGCAGCCACGGA	487
Db	43	-----	43
Qy	488	ATAAGGCCGAGGAAGAGAGAGAAACAAGATGGGGAAACACAGGTGCCATGTGAGG	547
Db	43	-----	43
Qy	548	GAAGAAACATCTCGGAACAAACCTTTCTACTTCCTCGTCAAGGGCGTTTAGCACCCGCTAC	607
Db	44	-----AlaGlnAsnAsnProTyrLeuPheArgSerAsnLysPheLeuThrLeuPhe	60
Qy	608	GGGAACCAAAACGTAGGATCCGGCTCTCCAGAGGTTTCACCAAAAGGTCAAGCAGCTTT	667
Db	61	LysAsnGlnHisGlySerLeuArgLeuGlnArgPheAsnGluAspThrGluLysLeu	80
Qy	668	CAGAATCTCCAGAAATCACCGTATTGTGCAGATCGAGCCCAACCTAACACTCTGTGCTTT	727
Db	81	GluAsnLeuArgAspTyrArgValLeuGluTyrCysSerLysProAsnThrLeuLeuLeu	100
Qy	728	CCAAGCACCGCTGATGCTGATACATCCTTGTGTTATCCAGCAGGGCAGCCACCGTGACC	787
Db	101	ProHisHisSerAspSerAspLeuLeuValLeuValLeuGluGlyGlnAlaLeuVal	120
Qy	788	GTAGCAAAATGGCAATACAGAAAGAGCTTTAACTTTGACGAGGCGCCATCAGATCAGAATC	847
Db	121	LeuValAsnProAspGlyArgAspThrTyrLysLeuAspGlnGlyAspAlaIleLysIle	140
Qy	848	CCATCCGGTTTCATTTCTTACATCTTGAACCGCCGATGACACCAACGAGACCTCAGAGTAGCT	907
Db	141	GluAlaGlyThrProPheTyrLeuIleAsnProAspAsnAsnGlnAsnLeuArgIleLeu	160
Qy	908	AAATCTCCATGCCCGTTAAACACACCGCGCAGTTTGTAGGATTTCTTCCCGGCGAGCAGC	967
Db	161	AsnPheAlaIleThrPheArgArgProGlyThrValGluAspPhePheLeuSerSerThr	180
Qy	968	CGAGACCAATCATCTTACGAGGCTTTCAGCAGGATACGTTGGAGGCGCGCTTCAAT	1027
Db	181	LysArgLeuProSerTyrLeuSerAlaPheSerLysAsnPheLeuGluAlaSerTyrAsp	200

```
QY 1028 GCGGAATTCATAGATACGAGGAGGTGCTGTAGAAAGAAATGCGAGGAGTGAGCAAG 1087
Db 1028 GCGGAATTCATAGATACGAGGAGGTGCTGTAGAAAGAAATGCGAGGAGTGAGCAAG 1087
QY 201 SerProTyrAspGluGluGlnThrLeuLeuGln----- 212
Db 201 SerProTyrAspGluGluGlnThrLeuLeuGln----- 212
QY 1088 GAGAGAGGCGAGGCGAGTACTCGGAGTAGTAGAACAATGAAGAGTGTATGTC 1147
Db 1088 GAGAGAGGCGAGGCGAGTACTCGGAGTAGTAGAACAATGAAGAGTGTATGTC 1147
QY 213 -----GluGluGlnGluGlyValileVal 220
Db 213 -----GluGluGlnGluGlyValileVal 220
QY 1148 AAAGTGTCAAAGGAGCAGCTTGAAGAACTTAAAGCAAGCTTAAATCCGTCTCAAGAAA 1207
Db 1148 AAAGTGTCAAAGGAGCAGCTTGAAGAACTTAAAGCAAGCTTAAATCCGTCTCAAGAAA 1207
QY 221 LysMetProLysAspGlnGluGlnLeuSerLysHisAlaGlnSerSerArgGly 240
Db 221 LysMetProLysAspGlnGluGlnLeuSerLysHisAlaGlnSerSerArgGly 240
QY 1208 GGCTCCGAGAGAGGAGATATACCAACCAATCACTTACGAGAGGCGAGCCGAT 1267
Db 1208 GGCTCCGAGAGAGGAGATATACCAACCAATCACTTACGAGAGGCGAGCCGAT 1267
QY 241 ThrLeuSerSerGln-----AspLysProPheAsnLeuArgSerArgAspProile 257
Db 241 ThrLeuSerSerGln-----AspLysProPheAsnLeuArgSerArgAspProile 257
QY 1268 CTTTCTCAACTTTGGGAAGTATTATGAGTGAAGCCAGCAGACAGAAAGACCCCGCTT 1327
Db 1268 CTTTCTCAACTTTGGGAAGTATTATGAGTGAAGCCAGCAGACAGAAAGACCCCGCTT 1327
QY 258 TyrSerAsnAsnTyrGlyLeuTyrGluThrProGlu-----LysAsnSerGlnLeu 276
Db 258 TyrSerAsnAsnTyrGlyLeuTyrGluThrProGlu-----LysAsnSerGlnLeu 276
QY 1328 CAGGACCTGGACATGCTCCTCCTGTAGATCAAGAGAGGCTTTGATGCTCCCA 1387
Db 1328 CAGGACCTGGACATGCTCCTCCTGTAGATCAAGAGAGGCTTTGATGCTCCCA 1387
QY 277 ArgAspLeuAspIleLeuLeuAsnCysLeuGlnMetAsnGluGlyAlaLeuPheValPro 296
Db 277 ArgAspLeuAspIleLeuLeuAsnCysLeuGlnMetAsnGluGlyAlaLeuPheValPro 296
QY 1388 CACTTCAACTCAAAGGCCATGCTTATCGTCTCTCAACAAAGAACTGGAAACCTTGAA 1447
Db 1388 CACTTCAACTCAAAGGCCATGCTTATCGTCTCTCAACAAAGAACTGGAAACCTTGAA 1447
QY 297 HisTyrAsnSerArgAlaThrValIleLeuValAlaAsnGluGlyArgAlaGluValGlu 316
Db 297 HisTyrAsnSerArgAlaThrValIleLeuValAlaAsnGluGlyArgAlaGluValGlu 316
QY 1448 CTCGTGGCTGTAAGAAAGAGCAACACAGAGGGGCGCGGGGAGAGAGGAGGAGCA 1507
Db 1448 CTCGTGGCTGTAAGAAAGAGCAACACAGAGGGGCGCGGGGAGAGAGGAGGAGCA 1507
QY 317 LeuValGlyLeuGluGlnGlnGln----- 326
Db 317 LeuValGlyLeuGluGlnGlnGln----- 326
QY 1508 GACGAAGAGAGGAGGA---AGTAAACAGAGAGTGGTGGTATGACAGCGAGGTTGAAG 1564
Db 1508 GACGAAGAGAGGAGGA---AGTAAACAGAGAGTGGTGGTATGACAGCGAGGTTGAAG 1564
QY 327 -----GlyLeuGluSerMetGlnLeuArgArgTyrAlaAlaThrLeuSer 341
Db 327 -----GlyLeuGluSerMetGlnLeuArgArgTyrAlaAlaThrLeuSer 341
QY 1565 GAAGCGCATGTGTTATCATGCGAGCAGCTCATCCAGTACCATCCATCAAGCTTCCCGAA 1624
Db 1565 GAAGCGCATGTGTTATCATGCGAGCAGCTCATCCAGTACCATCCATCAAGCTTCCCGAA 1624
QY 342 GluGlyAspIleLeuValIleProSerSerPheProValAlaLeuLysAlaLaSerAsp 361
Db 342 GluGlyAspIleLeuValIleProSerSerPheProValAlaLeuLysAlaLaSerAsp 361
QY 1625 CTCATCTGCTGCTGCTGCTATCAACGCTGAAACCAACACAGAACTTCTTCTGCGGT 1684
Db 1625 CTCATCTGCTGCTGCTGCTATCAACGCTGAAACCAACACAGAACTTCTTCTGCGGT 1684
QY 362 LeuAsnMetValGlyIleGlyValAsnAlaGluAsnAsnGluArgAsnPheLeuAlaGly 381
Db 362 LeuAsnMetValGlyIleGlyValAsnAlaGluAsnAsnGluArgAsnPheLeuAlaGly 381
QY 1685 GATAAGACATGTGATAGACGATAGAGAACGAGGAGGATGATGATTCCTCTGCGG 1744
Db 1685 GATAAGACATGTGATAGACGATAGAGAACGAGGAGGATGATGATTCCTCTGCGG 1744
QY 382 AsnLysGluAsnValIleArgGlnIleProArgGlnValSerAspLeuThrPheProGly 401
Db 382 AsnLysGluAsnValIleArgGlnIleProArgGlnValSerAspLeuThrPheProGly 401
QY 1745 TCGGCTGAACAGTTGAGAGCTCATCAAAACAGAGGAGGATCTCCTTGTGAGTGCT 1804
Db 1745 TCGGCTGAACAGTTGAGAGCTCATCAAAACAGAGGAGGATCTCCTTGTGAGTGCT 1804
QY 402 SerGlyGluGluValGluGluLeuLeuGluAsnGlnLysGluSerTyrPheValAspGly 421
Db 402 SerGlyGluGluValGluGluLeuLeuGluAsnGlnLysGluSerTyrPheValAspGly 421
QY 1805 CGTCTCAA 1813
Db 1805 CGTCTCAA 1813
QY 422 GlnProArg 424
Db 422 GlnProArg 424
RESULT 14
T44430
protein PVI00 [imported] - winter squash
C:Species: Cucurbita maxima (winter squash)
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C/Accession: T44430
R/Yamada, K.; Shimada, T.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I.
J. Biol. Chem. 274, 2563-2570, 1999
A>Title: Multiple functional proteins are produced by cleaving Asn-Gln bonds of a single
A/Reference number: 222767; MUID:99107919; PMID:9891029
A/Accession: T44430
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-810 <X>
A/Cross-references: UNIPROT:Q9ZWT3; EMBL:AB019195; NID:G3808061; PIDN:BAA34056.1; PID:93
Alignment Scores:
```

```
Pred. No.: 1-4e-55 Length: 810
Score: 907.50 Matches: 214
Percent Similarity: 54.65% Conservative: 121
Best Local Similarity: 34.91% Mismatches: 191
Query Match: 25.48% Indels: 87
DB: 2 Gaps: 20
US-10-728-051-1 (1-2032) x T44430 (1-810)
QY 188 CAACAGAACCCGAT---GACTTGAAGAAAAGGATGCGAGTCTCGTGCACCAAGCTC 244
Db 188 CAACAGAACCCGAT---GACTTGAAGAAAAGGATGCGAGTCTCGTGCACCAAGCTC 244
QY 244 GlnArgAspProAspTyrArgArgGluGlnLysArgArgGluGlnGlnGluArgArg 263
Db 244 GlnArgAspProAspTyrArgArgGluGlnLysArgArgGluGlnGlnGluArgArg 263
QY 245 GAGTATGATCTCTGTTGTTCTATGATCTCGAGGACACACTGGGACCAACCAACCT 304
Db 245 GAGTATGATCTCTGTTGTTCTATGATCTCGAGGACACACTGGGACCAACCAACCT 304
QY 264 GlnArgGluArgGlyGlyArgAspAspGluAspGlu-----AsnGlnArg 279
Db 264 GlnArgGluArgGlyGlyArgAspAspGluAspGlu-----AsnGlnArg 279
QY 305 TCCCTCCAGGAGGAGCGGACACGTCGCCGCCCAACCCGAGAC---TACGATGATGACCG 361
Db 305 TCCCTCCAGGAGGAGCGGACACGTCGCCGCCCAACCCGAGAC---TACGATGATGACCG 361
QY 280 AspProAspTyrArgArgGluGlnGluArgArgGluGlnGluArgArgArgGluArg 299
Db 280 AspProAspTyrArgArgGluGlnGluArgArgGluGlnGluArgArgArgGluArg 299
QY 362 CGTCAACCCGAGAGAGGAGGAGCCGATGGGACCCAGCTGGACCCGAGGAGGCTGAA 421
Db 362 CGTCAACCCGAGAGAGGAGGAGCCGATGGGACCCAGCTGGACCCGAGGAGGCTGAA 421
QY 300 GlnGlnGluArgArgGluGluHisArgGlyGlyArgValGluAspGluAsnGln 319
Db 300 GlnGlnGluArgArgGluGluHisArgGlyGlyArgValGluAspGluAsnGln 319
QY 422 AGAGAAGAGACTGGAGACCAACAGAGAGATTTGGGCGGACCAAGTCATCAGCAGCCA 481
Db 422 AGAGAAGAGACTGGAGACCAACAGAGAGATTTGGGCGGACCAAGTCATCAGCAGCCA 481
QY 320 ArgAspProAspTyrArgArgGluGlnGluArgArgGluGluGluGluGluArgGlu 339
Db 320 ArgAspProAspTyrArgArgGluGlnGluArgArgGluGluGluGluGluArgGlu 339
QY 482 CGGAAA---ATAAGGCCCGAGAGAGAGAGAGAACCAAGAGTGGGGAACACAGGTAGC 538
Db 482 CGGAAA---ATAAGGCCCGAGAGAGAGAGAGAACCAAGAGTGGGGAACACAGGTAGC 538
QY 340 ArgGluTyrGluArgGluHisGlyArgArgGlyArgGluGluGlnArgSerArgGlu 359
Db 340 ArgGluTyrGluArgGluHisGlyArgArgGlyArgGluGluGlnArgSerArgGlu 359
QY 539 CATGTGAGG----- 547
Db 539 CATGTGAGG----- 547
QY 360 GlnArgArgArgHisGluArgGlnHisGlyArgSerArgValAsnGlnValAlaIle 379
Db 360 GlnArgArgArgHisGluArgGlnHisGlyArgSerArgValAsnGlnValAlaIle 379
QY 548 -----GAAGAAACATCTCGGAACCAACCTTCTACTTCCCGTCAAGGCGGTTAGC 598
Db 548 -----GAAGAAACATCTCGGAACCAACCTTCTACTTCCCGTCAAGGCGGTTAGC 598
QY 380 ArgArgThrGluGlnGluGlnSerAsnAsnProTyrTyrPheGlnGluGlnArgPheGln 399
Db 380 ArgArgThrGluGlnGluGlnSerAsnAsnProTyrTyrPheGlnGluGlnArgPheGln 399
QY 599 ACCCGCTAGGGAACCAAAAACGTTAGGATCCGGTCTGCGAGAGGTTTGAACCAAGGTCA 658
Db 599 ACCCGCTAGGGAACCAAAAACGTTAGGATCCGGTCTGCGAGAGGTTTGAACCAAGGTCA 658
QY 400 SerArgTyrArgSerAspGluGlyHisTyrArgValLeuGluArgPheSerGluArgSer 419
Db 400 SerArgTyrArgSerAspGluGlyHisTyrArgValLeuGluArgPheSerGluArgSer 419
QY 659 AGGCGATTTCAAGATCTCCAGAAATCCCGTATTGCGAGATCGAGGCCAAACCTAACACT 718
Db 659 AGGCGATTTCAAGATCTCCAGAAATCCCGTATTGCGAGATCGAGGCCAAACCTAACACT 718
QY 420 GlnLeuLeuLysGlyIleLysAsnGlnArgLeuAlaLeuLeuGluAlaArgProHisThr 439
Db 420 GlnLeuLeuLysGlyIleLysAsnGlnArgLeuAlaLeuLeuGluAlaArgProHisThr 439
QY 719 CTTGTTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTCTTCCAGCAAGGCGCAAGCC 778
Db 719 CTTGTTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTCTTCCAGCAAGGCGCAAGCC 778
QY 440 PheIleValProHisHisLeuAspAlaGluCysValLeuLeuValValArgGlyArgAla 459
Db 440 PheIleValProHisHisLeuAspAlaGluCysValLeuLeuValValArgGlyArgAla 459
QY 779 ACCGTGACCGTAGCA-----AATGGCAATACAGAAAGAGAGGCTTTAATCTTGAC 826
Db 779 ACCGTGACCGTAGCA-----AATGGCAATACAGAAAGAGAGGCTTTAATCTTGAC 826
QY 460 ThrIleThrThrValValGlnGluLysArgGluThrArgLysGluSerTyrAsnValGlu 479
Db 460 ThrIleThrThrValValGlnGluLysArgGluThrArgLysGluSerTyrAsnValGlu 479
QY 827 GAGGCGCATGCACTCAGAAATCCCGTGTTCATTTCTTCTACATCTTGAACCGCATGAC 886
Db 827 GAGGCGCATGCACTCAGAAATCCCGTGTTCATTTCTTCTACATCTTGAACCGCATGAC 886
QY 480 SerGlyAspValMetThrIleProAlaGlyThrThrLeuTyrLeuAlaAsnGln---Glu 498
Db 480 SerGlyAspValMetThrIleProAlaGlyThrThrLeuTyrLeuAlaAsnGln---Glu 498
QY 887 AACCAAGACCTCAGAGTAGCTTAAATCTCCATGCCCGTTAAACACACCCGCGCAGTTTGAG 946
Db 887 AACCAAGACCTCAGAGTAGCTTAAATCTCCATGCCCGTTAAACACACCCGCGCAGTTTGAG 946
QY 499 AsnGluAspLeuGlnIleValLysLeuValGlnProValAsnAsnProGlyGluPheLys 518
Db 499 AsnGluAspLeuGlnIleValLysLeuValGlnProValAsnAsnProGlyGluPheLys 518
QY 947 GATTCTTCCCGGCGAGCGGAGCCGAGACCAATCATCTTCTTCTGAGGCGCTTACGAGGAAT 1006
Db 947 GATTCTTCCCGGCGAGCGGAGCCGAGACCAATCATCTTCTTCTGAGGCGCTTACGAGGAAT 1006
QY 519 AspTyrLeuSerAlaGlyGlyGluSerGln---AlaTyrTyrSerValPheSerAsnAsp 537
Db 519 AspTyrLeuSerAlaGlyGlyGluSerGln---AlaTyrTyrSerValPheSerAsnAsp 537
QY 1007 ACGTTGGAGCGCGCTTCAATGCGGAATTCATGAGATACGAGGCGTGTGTTAGAGAG 1066
Db 1007 ACGTTGGAGCGCGCTTCAATGCGGAATTCATGAGATACGAGGCGTGTGTTAGAGAG 1066
QY 538 ValLeuGluAlaLeuAsnIleProArgAspLysLeuGluArgIlePhe----- 554
Db 538 ValLeuGluAlaLeuAsnIleProArgAspLysLeuGluArgIlePhe----- 554
```

A:Molecule type: DNA  
A:Residues: 1-605 <CHL>  
A:Cross-references: UNIPROT:P09799  
C:Superfamily: glycinin  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-605/Product: alpha-globulin type A #status predicted <MAT>

Alignment Scores:  
Pred. No.: 8,2e-51 Length: 605  
Score: 839.00 Matches: 216  
Percent Similarity: 50.86% Conservative: 108  
Best Local Similarity: 33.91% Mismatches: 223  
Query Match: 23.55% Indels: 90  
DB: 2 Gaps: 21

US-10-728-051-1 (1-2032) x S06398 (1-605)

Qy	65	TC	TC	CA	CT	GAT	GCT	GT	GT	CT	AG	GA	CT	CT	GT	CT	GC	TT	CAG	TT	CT	GC	CA	CG	CA	TGC	124
Db	6	Ser	Val	Phe	Val	Val	Leu	Leu	Phe	Ser	Leu	Phe	Leu	Ser	Phe	Gly	Leu	Leu	Cys	Ser	Ala	25					
Qy	125	AAG	TC	AT	CA	CT	TAC	CA	GA	AA	AA	CA	AG	AA	CC	CT	CG	CC	CAG	AGG	-----	TGC	175				
Db	26	Lys	Asp	Phe	Pro	Gly	Arg	Arg	Ser	Glu	Asp	Asp	Pro	-----	Gln	Gln	Arg	Tyr	Glu	Asp	Cys	44					
Qy	176	CT	CA	GAG	TGT	CA	CA	CA	AGG	ACCG	---	GAT	GAC	TT	TGA	AG	CA	AA	AGG	CA	TGC	232					
Db	45	Arg	Lys	Arg	Cys	Gln	Leu	Glu	Thr	Arg	Gly	Gln	Thr	Glu	Gln	Asp	Lys	Cys	Glu	Asp	Arg	64					
Qy	233	TGC	---	ACC	AAG	CT	CG	AG	-----	-----	-----	TAT	GAT	CT	CT	CT	GT	GT	GT	CT	AT	268					
Db	65	Ser	Glu	Thr	Gln	Leu	Lys	Glu	Gln	Gln	Arg	Asp	Gly	Glu	Asp	Pro	Gln	Arg	Arg	Tyr	84						
Qy	269	---	GAT	CT	CG	AG	GAC	ACT	TGG	CACC	CA	CA	CG	TT	CC	CT	C	CAG	GG	-----	316						
Db	85	Gln	Asp	Cys	Arg	Gln	His	Cys	Gln	Gln	Gln	Glu	Arg	Leu	Arg	Pro	His	Cys	Glu	Gln	104						
Qy	317	---	---	---	---	GAG	CG	GAC	ACT	TGG	CG	CG	CA	CC	CGG	AG	ACT	ACG	AT	GAT	GAC	358					
Db	105	Ser	Cys	Arg	Glu	Gln	Tyr	Glu	Lys	Gln	Gln	Gln	Gln	Pro	-----	Asp	119										
Qy	359	CG	CG	TC	CA	CC	CG	AG	AG	AGG	AGG	CG	CG	AT	TGG	GAC	CAG	CT	GAC	CG	-----	409					
Db	120	Lys	Gln	Phe	Lys	Glu	Cys	Gln	Arg	Cys	Gln	Trp	Gln	Glu	Gln	Arg	Pro	Glu	Arg	Lys	139						
Qy	410	---	---	---	---	AGG	AG	CG	TGA	AG	AG	AA	CA	AG	ACT	TGG	CA	CA	CA	CA	448						
Db	140	Gln	Gln	Cys	Val	Lys	Glu	Cys	Arg	Ser	Glu	Gln	Tyr	Gln	Glu	Asp	Pro	Tyr	Lys	Gly	159						
Qy	449	GA	AG	AT	---	TGG	AG	CGG	CA	CC	AA	GT	CT	AT	CAG	CA	CG	CA	CG	AA	505						
Db	160	Glu	Asn	Lys	Trp	Arg	Glu	Glu	Glu	Glu	---	---	---	---	---	---	---	---	---	---	172						
Qy	506	GA	AG	GA	AA	CA	AG	TGG	GG	GA	CA	CC	AG	TAG	CC	AT	GTG	AGG	GA	AA	565						
Db	173	Glu	Gly	Glu	Gln	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	180						
Qy	566	AA	CC	TT	CT	TAC	TT	CC	CG	T	CA	AG	CG	GT	TT	TAG	CA	CC	CG	CT	625						
Db	181	Asn	Pro	Tyr	Tyr	Phe	His	Arg	Arg	Ser	Phe	Gln	Arg	Phe	Arg	Glu	Glu	His	Gly	Asn	200						
Qy	626	AT	CC	GG	CT	CT	CG	CA	AG	GT	TT	GA	CA	AA	GG	CT	CA	AG	CG	AT	685						
Db	201	Phe	Arg	Val	Leu	Gln	Arg	Phe	Ala	Asp	Lys	His	Ile	Leu	Arg	Gly	Ile	Asn	Glu	Phe	220						
Qy</																											

007  
119475

Qy	626	ATCCGGTCTCGAGAGGTTTGACCAAGGAGTTCAGAAATCTTCAGAAATCCAC	685
Qy	626	ATCCGGTCTCGAGAGGTTTGACCAAGGAGTTCAGAAATCTTCAGAAATCCAC	685
Db	201	PheArgValLeuGlnArgPheAlaAspLysHisHisLeuLeuArgGlyIleAsnGluPhe	220
Qy	686	CGTATTGTGCAGATCGAGGCCAAACCTAACACTCTGTCTTCTCCAGACAGCGTGATGCT	745
Db	221	ArgIleAlaIleLeuGluAlaAsnProAsnThrPheValLeuProHisHisCysAspAla	240
Qy	746	GATAACATCTTGTATTCCAGCAAGGCCAGCCACCGTAGCCGATAGCAAAATGCCAATAAC	805
Db	241	GluLysIleTyrValValThrAsnGlyArgGlyThrValThrPheValThrHisGluAsn	260



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 8, 2005, 03:45:11 ; Search time 343.105 Seconds  
(without alignments)  
6065.459 Million cell updates/sec

Title: US-10-728-051-1

Perfect score: 3562  
Sequence: 1 aataacatataattcattc.....cgtttggtggtttcttc 2032

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1612378 segs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2 1/USPTO.spool\_p/US10728051/runat\_07042005\_125045\_14049/app.query.fasta\_1.4757  
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0  
-UNITS=Dits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USR=US10728051@CGN 1.1 874 @runat\_07042005\_125045\_14049 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	92.3	626	1 AL12_ARAHY	P43238 arachis hyp
2	3041	85.4	614	1 AL11_ARAHY	P43237 arachis hyp
3	2887.5	81.1	580	2 Q6PSU3	Q6psu3 arachis hyp
4	2172.5	61.0	428	2 Q6PSU4	Q6psu4 arachis hyp
5	1495.5	42.0	299	2 Q6PSU5	Q6psu5 arachis hyp
6	1435	40.3	303	2 Q6PSU6	Q6psu6 arachis hyp
7	1302	36.6	621	2 Q7XHT2	Q7xht2 glycine max
8	1290	36.2	533	2 Q6EBG1	Q6ebg1 lupinus alb
9	1284	36.0	621	2 Q948T0	Q948t0 glycine max
10	1282	36.0	605	2 Q94LX2	Q94lx2 glycine max
11	1281	36.0	605	1 GLCA_SOYBN	P13916 glycine max
12	1267	35.6	571	1 CVCA_PEA	P13915 pisum sativ
13	1265.5	35.5	623	2 Q948X9	Q948x9 glycine max
14	1258.5	35.3	545	2 Q41674	Q41674 vicia narbo
15	1251	35.1	613	2 Q9M3X6	Q9m3x6 pisum sativ
16	1251	35.1	639	1 GLCX_SOYBN	P11827 glycine max

17	1190	33.4	559	2 Q9FP29	Q9fzp9 glycine max
18	1183.5	33.2	418	2 Q84U11	Q84u11 lens culina
19	1180.5	33.1	543	2 Q22120	Q22120 glycine max
20	1178.5	33.1	459	1 VCLC_PEA	P13918 pisum sativ
21	1168	32.8	415	2 Q84U10	Q84u10 lens culina
22	1165	32.7	416	2 Q22121	Q22121 glycine max
23	1156	32.5	439	1 GLCB_SOYBN	P25974 glycine max
24	1152	32.3	439	2 Q93VL9	Q93vl9 glycine max
25	1148	32.2	438	2 Q43626	Q43626 pisum sativ
26	1145	32.1	415	2 Q702P1	Q702p1 pisum sativ
27	1135	31.9	415	2 Q702P0	Q702p0 pisum sativ
28	1123.5	31.5	453	2 Q41677	Q41677 vicia narbo
29	1108.5	31.1	483	1 VCL_VICFA	P08438 vicia faba
30	1082	30.4	396	2 Q84UB3	Q84ub3 glycine max
31	1068	30.0	410	1 VCLB_PEA	P02854 pisum sativ
32	1030	28.9	518	2 Q9M3X8	Q9m3x8 lens culina
33	962.5	27.0	445	1 CANA_CANEN	P50477 canavalia e
34	959.5	26.9	445	1 CANA_CANGL	P10562 canavalia g
35	910	25.5	533	2 Q9SEW4	Q9sew4 juglans reg
36	907.5	25.5	810	2 Q9ZWI3	Q9zwi3 cucurbita m
37	880	24.7	481	2 Q7Y1C1	Q7y1c1 juglans nig
38	865.5	24.3	666	2 Q9SPL4	Q9spl4 macadamia i
39	863.5	24.2	625	2 Q9SPL3	Q9spl3 macadamia i
40	849.5	23.8	666	2 Q9SPL5	Q9spl5 macadamia i
41	839	23.6	605	1 VCLA_GOSHI	P09799 gossypium h
42	825.5	23.2	436	1 PHSA_PHAUV	P07219 phaseolus v
43	806.5	22.6	430	2 Q41115	Q41115 phaseolus v
44	805.5	22.6	430	2 Q43633	Q43633 phaseolus v
45	804	22.6	423	1 PHS2_PHALU	Q43617 phaseolus l

#### ALIGNMENTS

RESULT 1

AL12\_ARAHY  
ID AL12\_ARAHY STANDARD; PRT; 626 AA.

AC P43238;  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Allergen Ara h 1, clone P41B precursor (Ara h I).  
OS Arachis hypogaea (Peanut).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
OC Arachis.  
OX NCBI\_TaxID=3818;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Florunner;  
RX MEDLINE=96013631; PubMed=7560062;  
RA Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;  
RT "Recombinant peanut allergen Ara h I expression and IGE binding in  
RT patients with peanut hypersensitivity.";  
RL J. Clin. Invest. 96:1715-1721(1995).  
RN [2]  
RP CARBOHYDRATE-LINKAGE SITE ASN-521.  
RX MEDLINE=20455243; PubMed=10998264; DOI=10.1006/abio.2000.4737;  
RA Kolarich D., Altmann F.;  
RT "N-glycan analysis by matrix-assisted laser desorption/ionization mass  
RT spectrometry of electrophoretically separated nonmammalian proteins:  
RT application to peanut allergen Ara h I and olive pollen allergen Ole e  
RT 1.";  
RL Anal. Biochem. 285:64-75(2000).  
CC -I- ALLERGEN. Causes an allergic reaction in human.  
CC -I- SIMILARITY: Belongs to the 7S seed storage protein family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-eb.ch/announce/

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L34402; AAB00861.1; -.
DR HSSP; P25974; 1IPJ.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmC_Like_cupin.
DR Pfam; PF00190; Cupin; 2.
KW Allergen; Glycoprotein; Signal.
FT SIGNAL 1 25 potential.
FT CHAIN 26 626 Allergen Ara h 1, clone P41B.
FT CARBOHYD 521 521 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 626 AA; 71345 MW; 1A6BBE41490D0E3 CRC64;

Alignment Scores:
Pred. No.: 2,34e-207 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 1 Gaps: 0

US-10-728-051-1 (1-2032) x AL12_ABAHY (1-626)
QY 50 ATGAGAGGAGGTTTCTCCTCAGTATGCTGTTCTAGGATCCTTGTCTCGGTTGAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATCCCAAGTCATCACCTTACCAGAAGAAACAGAACCCCTGCGCCAG 169
DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCTCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGACCAAGCTCAGATATGATCTCTGTTGTCTATGATCTCTCGAGACACATGCG 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACAACCAACGTTCCCTCTCAGGGAGCGACACGTCGGCGCCCAACCCGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGGAGGCGGATGGGACCGACGCTGGACCG 409
DB 101 AspAspAspArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
QY 410 AGGAGGCTGAAGAGAGAGAGACTGGAGACAACCAAGAGAAATGGAGGCGACCAAGT 469
DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProser 140
QY 470 CATCAGCAGCCACGGAATAAGCCCGAGGAAGAGAGAGCAACAGAGTGGGGAACA 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnLutProGlyThr 160
QY 530 CCAAGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTTCTACTTCCCGTCAAG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGGAGTTTCAAGATCTCCAGAATCACCGTATTGTGCAGATCCGAGGCCAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTTGTCTTCCCAAGCAGCTGATGCTGATTAACATCTTGTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGln 240
QY 770 GGCAAGCCACCGTGACCGTAGCAATGGCAATTAACAGAAAGCTTTAATCTTGACGAG 829

```

---

```

241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
830 GGCCATGCACCTCAGAAATCCCATCCCGTTTCATTTCTACATCTTGAACCGCATGACAA 889
261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
890 CAGAACTTCAGAGTAGCTAAATCTCCATGCCGTGTAACACACCCGCGCAGTTGAGGAT 949
281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
950 TTCCTCCCGGCGAGCAGCCGAGACCAATCATCTTCTTCCAGGGCTTCAGCAGGATAG 1009
301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
1010 TTGGAGCGCGCTTCAATCGGAAATTCATAGATACGAGGGTGTGTTTAGAAGAGAT 1069
321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
1070 GCAGGAGGTGACGAAGAGAGAGAGCGGAGAGCGATGAGTACTCGGAGTAGTAGAAG 1129
341 AlaGlyGlyGluGlnGluArgGlyGlnArgTyrSerThrArgSerSerGluAsn 360
1130 AATGAGGAGTATAGTCAAGTCTCAAGGAGCGAGCTTGAAGAACTTACTTAAGCAGCT 1189
361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
1190 AATCGCTCTCAAGAAAGCGTCCGAAGAGAGGAGATATCAACCAACCAATCAACTTG 1249
381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
1250 AGAGAAGCGAGCCCGATCTTTCTAACTTTGGGAAGTTATTTGAGTGAAGCCAGAC 1309
401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
1310 AGAGAAGCCCGAGCTTCAGGACCTGACATGATGCTCCTCCTGTCAGATCAAGAA 1369
421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 440
1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCGCATGTTATCGTCTCGTCAACAA 1429
441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
1430 GGAATCGGAACTTGAATCTGCTGCTGTAAGAAAGAGCAACACAGAGGGAGCGCGG 1489
461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 480
1490 GAAGAGAGGAGGAGCAAGACGAGAGAGAGGAGGAGTAACAGAGAGTGCCTAGGTAC 1549
481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
1550 ACAGCGAGGTTCAAGGAAGCGATGTTTCATCATGCGCAGCAGCTCATCCAGTAGCCATC 1609
501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
1610 AACGTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCAAGA 1669
521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
1670 ATCTTCTTCAGGTCATAGGACATGATGTAGACCATAGAGAGCAAGCAAGGAGGAT 1729
541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
1730 TTAGCATTCCTCGGTCGGGTGAACAGTTGAGAGCTCATCAAAACCAAGAGGAATCT 1789
561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
1790 CACTTTGTGAGTCTGCTCCTCAATCTCAATCTCGTCTCGTCTCTCTGAGAAAGAG 1849
581 HisPheValSerAlaArgProGlnSerGlnSerProSerSerProGluLysGlu 600
1850 TCTCTCGAAGAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGTCCACTCTCTTCA 1909

```



Db	601	SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
Qy	1910	ATTTTGAAGCTTTTAAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
RESULT 2			
AL11_ARAHY			
ID	AL11_ARAHY	STANDARD;	PRT; 614 AA.
AC	P43237;		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	10-NOV-1995 (Rel. 32, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Allergen Ara h 1, clone P17 precursor (Ara h I).		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurooids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis		
OX	NCBI_TaxID=3818;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Florunner;		
RX	MDLINE=96013631; PubMed=7560062;		
RA	Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;		
RT	"Recombinant peanut allergen Ara h I expression and IgE binding in		
RT	patients with peanut hypersensitivity.";		
RL	J. Clin. Invest. 96:1715-1721(1995).		
RN	[2]		
RP	CARBOHYDRATE-LINKAGE SITE ASN-516.		
RX	MDLINE=20455243; PubMed=1098264; DOI=10.1006/abio.2000.4737;		
RA	Kolarich D., Altmann F.;		
RT	"N-glycan analysis by matrix-assisted laser desorption/ionization mass		
RT	spectrometry of electrophoretically separated nonmammalian proteins:		
RT	application to peanut allergen Ara h 1 and olive pollen allergen Ole e		
RT	1.";		
RL	Anal. Biochem. 285:64-75(2000).		
CC	-/- ALLERGEN: Causes an allergic reaction in human.		
CC	-/- SIMILARITY: Belongs to the 7S seed storage protein family.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	ENBL; L38853; AAA60336.1; -.		
DR	HSSP; P25974; IIPJ.		
DR	GlycoSuiteDB; P43237; -.		
DR	InterPro; IPR006045; Cupin.		
DR	InterPro; IPR007113; Cupin region.		
DR	InterPro; IPR011051; RmlC_Like_cupin.		
DR	Pfam; PF00130; Cupin; 2.		
KW	Allergen; Glycoprotein; Signal.		
FT	SIGNAL	25	Potential.
FT	CHAIN	1 26 614	Allergen Ara h 1, clone P17.
FT	CARBOHYD	516 516	N-linked (GlcNAc...).
FT			/FTId-CAR_000218.
SQ	SEQUENCE	614 AA; 70283 MW; 1DDACF217EBC5F31 CRC64;	
Alignment Scores:			
Pred. No.:	2.9e-191	Length:	614
Score:	3041.00	Matches:	595
Percent Similarity:	96.03%	Conservative:	9
Best Local Similarity:	94.59%	Mismatches:	7
Query Match:	85.37%	Indels:	18
DB:	1	Gaps:	7
US-10-728-051-1 (1-2032) x AL11_ARAHY (1-614)			
Qy	50	ATGAGAGGAGGGTTTCTCCATGCTGCTAGGGATCCCTTGTCTTCAGTT	109

Db	1	MetArgIArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal	20
Qy	110	TTGTCAACGATGCCAAGTCATCACCTTACCAAGAAAAACAGAGAACCCCTGCGCCAG	169
Db	21	SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln	38
Qy	170	AGGTGCTCCAGATTGTCTCAACAGGAACCGGATGACTTGAAGCAAAAGGATGCGAGTCT	229
Db	39	ArgCysLeuGlnSerCysGlnGluProAspAspLeuLysGlnIleAlaCysGluSer	58
Qy	230	CGCTGCACCAAGCTCCAGTATGATCTCTGTTGTCTATGATCTCCGAGGACACACATGCC	289
Db	59	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly	74
Qy	290	ACCACCAACCAACGTTCCCTCCAGGGGAGCGACACGTCGCGCCCAACCCGGAGATAC	349
Db	75	AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	94
Qy	350	GATGATGACCGCGCTCAACCCCAAGAGAGAGAGGCGCGATGGGGACCCAGCTGGACCG	409
Db	95	AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGluPro	114
Qy	410	AGGGCGCTGAAAGAGAGAGGACTCGAGACACCAAGAGAGAGATTCGAGCGCACCAAGT	469
Db	115	ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgProSer	134
Qy	470	CATCAGCAGCCACGGAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529
Db	135	HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTyrGlyThr	154
Qy	530	CAAGTAGCATGTGTGAGGAGAGAAACATCTCGGAACAAACCTTCTTCTTCCGCTCAAGG	589
Db	155	ProGlySerGluValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg	174
Qy	590	CGGTTTAGCACCCGCTACCGGACCAAAACGGTAGGATCCGGCTCTGCGAGAGTTTGAC	649
Db	175	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	194
Qy	650	CAAAGTGCAAGGAGTTTCCAGAAATCCAGAAATCCAGAAATCCAGAAATCCAGAAAT	709
Db	195	GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg	214
Qy	710	CCTAACACTCTTGTCTTCCCAAGCAGCTGATGCTGATAAACATCTTGTATCCAGCAA	769
Db	215	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	234
Qy	770	GGCAGACCCACCGTACCGCTAGCAATGGCAATGCAATGCAATGCAATGCAATGCAAT	829
Db	235	GlyGlnAlaThrValThrValAlaAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	254
Qy	830	GGCCATGCACTCAGAAATCCCATCCGCTTTTTCATTTCTTCTTCTTCTTCTTCTTCTT	889
Db	255	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	274
Qy	890	CAGAACCTCAGAGTAGCTAAATCTCCATGCGCGCTTAAACACACCCGCGCTTGAAGAT	949
Db	275	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	294
Qy	950	TTCTTCCCGGAGCAGCGGACCAATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1009
Db	295	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	314
Qy	1010	TTGGAGCGCGCTTCAATCGGAATTCATAGATACGAGGAGGCTGCTGTGTAGAGAGAT	1069
Db	315	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAlaAsn	334
Qy	1070	GCAGGAGGTGAGCAAG	1129
Db	335	AlaGlyGlyGluGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp---	353
Qy	1130	AATGAGGAGTGTAGTCAAGTGTCAAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAG	1189



```
Qy 890 CAGAACCTCAGATGAGTAAATCTCCATGCCCTTAAACACACCCGCCGCGATTGAGGAT 949
Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
Qy 950 TTCTTCCCGCGACGCCGAGACCAATCTCTACCTTGCAGGGCTTTCAGCAGGAATACG 1009
Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
Qy 1010 TTGAGCCGCCCTTCAATCGGAATTCATCAGATACGGAGGGTGTGTTAGAGAGAAT 1069
Db 315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 334
Qy 1070 GCAGGAGCTGACGAGAGAGAGAGCGGCAGAGCGATGGAGTACTCCGAGTAGTGAGAAC 1129
Db 335 AlaGlyGlyGlnGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
Qy 1130 AATGAAGGAGTGTAGTCAAGTCTCAAGAGGACGCTTGAAGCACTTACTTAAGCAGCT 1189
Db 354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373
Qy 1190 AAATCCGCTCTCAAGAAGGCTCCGAAGAAGAGGAGATATCACCAACCCCAATCAACTTG 1249
Db 374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
Qy 1250 AGAAGCGGAGCCGATCTTTCTTAAACAATTTGGGAAGTATTATTGAGGTGAGCCAGAC 1309
Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
Qy 1310 AAGAAGAACCCCGCTTCAGGACTCGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
Qy 1370 GAGCTTTGATGTCTCCACACTTCAACTCAAGGCCATGTTATCGTCGTCGCAACAAA 1429
Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452
Qy 1430 GGAACCTGGAACCTTGAACTCGTGGCTGTGAAGAAAGAGCAACAAACAGAGGGAGCGCG 1489
Db 453 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 472
Qy 1490 GAA-----GAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1540
Db 473 GluGlnGluTrpGluGluGluGluGluGluGluGluGluGluGluGluGluGluVal 492
Qy 1541 CGTAGGTACACAGCGAGGTTGAAGGAAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCA 1600
Db 493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512
Qy 1601 GTAGCCATCAACGCTTCCCGAACTCCATCTGCTTGGCTTGGCTTATCAACGCTGAAGAAC 1660
Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
Qy 1661 AACACAGAACTCTTCTTGCAGGTGATAAGACAAATGTATAGACCATAGAGAGAGAGAGAA 1720
Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
Qy 1721 GCGAAGGATTAGCATTTCCCTCGGTCGGGTGCAACAACTGTAGAGAGCTCATCAAAAACAG 1780
Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
Qy 1781 AAGGAATCTCACTTTGTGAGTGCT 1804
Db 573 ArgGluSerHisPheValSerAla 580
```

## RESULT 4

```
Q6PSU4 ID Q6PSU4 PRELIMINARY; PRT; 428 AA.
AC Q6PSU4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Conarachin (Fragment).
```

OS Arachis hypogaea (Peanut).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
OC Arachis.  
OX NCBI\_TaxID=3818;  
RN [1]\_TaxID=3818;  
RP SEQUENCE FROM N.A.  
RA Wang L., Liao B., Li H., Yan Y., Lin X., Huang S.;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY581851; AAT00596.1; -  
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
DR InterPro; IPR006045; Cupin.  
DR InterPro; IPR007113; Cupin\_region.  
DR InterPro; IPR011051; RnLC\_like\_cupin.  
DR Pfam; PF00190; Cupin; 2.  
FT NON TER 1  
SQ SEQUENCE 428 AA; 48094 MW; F1A32DABC33AD06F CRC64;

Alignment Scores:  
Pred. No.: 3,02e-134 Length: 428  
Score: 2172.50 Matches: 427  
Percent Similarity: 99.77% Conservative: 0  
Best Local Similarity: 99.77% Mismatches: 0  
Query Match: 60.99% Indels: 1  
DB: 2 Gaps: 1

US-10-728-051-1 (1-2032) x Q6PSU4 (1-428)

```
Qy 644 TTTGACCAAGGTCAGGCGAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAG 703
Db 2 PheAspGlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGlu 21
Qy 704 GCCAAACCTAAACACTCTTGTCTTCCCAAGCAGCTGATCTCTATAACATCTTGTATTATC 763
Db 22 AlalysProAsnThrLeuValLeuProLysHisAlaAspAlaAsnIleLeuValIle 41
Qy 764 CAGCAAGGGCAACCCGCTGACCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAACTTT 823
Db 42 GlnGlnGlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeu 61
Qy 824 GAGCAGGGCCATCAGCTCAGAAATCCCATCCGTTTCAATCTCTTACATCTTGTGACCGCAT 883
Db 62 AspGluGlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHis 81
Qy 884 GACAACCAAGCACTCAGAGTAGCTAAATCTCCATGCCCTTAAACACACCCGCGCAGTTT 943
Db 82 AspAsnGlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPhe 101
Qy 944 GAGGATTTCTTCCCGGAGCAGCCGAGACCAATCATCTTCTTACGAGGCTTCAGCAGG 1003
Db 102 GluAspPhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArg 121
Qy 1004 AATACCTTGGAGGCCCTTCAATCGGAATTCATCAGATACGAGGGTGTGTTAGAA 1063
Db 122 AsnThrLeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGlu 141
Qy 1064 GAGAATGTCAGGAGCTGAGCAAGAGAGAGAGAGAGGCGCAGAGCTGAGTAGTCTCGAGTAGT 1123
Db 142 GluAsnAlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerSer 161
Qy 1124 GAGAAACAATGAAGAGGTGATAGTCAAGTCTCAAGAGGACGCTTGAAGAACTTACTTAAG 1183
Db 162 GluAsnAsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLys 181
Qy 1184 CAGCTTAATCTCTCTCAAGAAGGCTCCGAAGAAGAGGAGGAGATATCACCAACCCCAATC 1243
Db 182 HisAlaLysSerValSerLysLysGlySer---GluGluGlyAspIleThrAsnProIle 200
Qy 1244 AACTTGAGAGAGAGCGAGCCGCGATCTTTCTTAAACAATTTGGGAAGTATTATTGAGGTGAAG 1303
Db 201 AsnLeuArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLys 220
```

```
QY 1304 CCAGACAAGAGAACCCAGCTTCAGGACCTGAGCATGATGCTACCTGTGTAGAGATC 1363
DB 221 ProAspLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIle 240
QY 1364 AAAGAAGAGCTTTGATGCTCCACACTTCAACTCAAGCCATGTTATCGTCTGCTC 1423
DB 241 LysGluGlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValVal 260
QY 1424 AACAAAGGACTGGAACCTTGAACCTCGCTGCTGTAAGAAAGAGCAACAAACAGAGGGA 1483
DB 261 AsnLysGlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGly 280
QY 1484 CGCGGGGAGAAAGAGGAGGACGAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1543
DB 281 ArgArgGluGluGluAspGluAspGluGluGluGluGluGluGluGluGluGluGlu 300
QY 1544 AGGTACACAGCGAGGTTGAAGAGAGGAGGATGTTTATCATGTCAGCAGCTCATCCAGTA 1603
DB 301 ArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProVal 320
QY 1604 GCCATCAACGCTTCCCTCCGAACTCCATCTGCTTGGCTTGGTATCAACGCTGAAAAACAAC 1663
DB 321 AlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsn 340
QY 1664 CACAGATCTTCTCTGAGGTGATAGGACAATGTGATAGACAGATAGAGAGCAAGCG 1723
DB 341 HisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAla 360
QY 1724 AAGATTAGCATTCCTCGGTGGTGAACAAGTTGAGAGAGTCAATCAAAACCAAGAG 1783
DB 361 LysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLys 380
QY 1784 GAATCTCACTTGTGAGTGTGCTGCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCT 1843
DB 381 GluSerHisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGlu 400
QY 1844 AAAGAGTCTCTGAGAAGAGGATCAAGAGAGAGAAACCAAGAGAGAGAGAGAGAGAGAG 1903
DB 401 LysGluSerProGluLysGluAspGlnGluGluGluGluGluGluGluGluGluGluGlu 420
QY 1904 CTTTCAATTTTGAAGGCTTTTAACT 1927
DB 421 LeuSerIleLeuLysAlaPheAsn 428
RESULT 5
Q6PSUS PRELIMINARY; PRT; 299 AA.
ID Q6PSUS AC Q6PSUS;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Conarachin (fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP SEQUENCE FROM N.A.
RA Wang L., Liao B., Li H., Yan Y., Lin X., Huang S.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY581850; AAT00595.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 1.
FT NON_TER 1
SQ SEQUENCE 299 AA; 33604 MW; 594C3AB0C2FD49EB CRC64;
```

Alignment Scores:

Pred. No.: 8.28e-90 Length: 299

RESULT 6

```
Score: 1495.50 Matches: 296
Percent Similarity: 97.39% Conservative: 2
Best Local Similarity: 96.73% Mismatches: 1
Query Match: 41.98% Indels: 7
DB: 2 Gaps: 1
US-10-728-051-1 (1-2032) x Q6PSU5 (1-299)
QY 1010 TTGGAGGCGCGCTTCAATGCGGAATTCATGAGATACGAGGGTCTCTTTAGAGAGAAAT 1069
DB 1 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 20
QY 1070 GCAGAGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1129
DB 21 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 40
QY 1130 AATGAAGAGAGTATAGTCAAAAGTGTCAAGAGAGACGCTTGAAGAACTTACTAAGCACGCT 1189
DB 41 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 60
QY 1190 AAATCCGTCTCAAGAAAGGCTCCGAAGAGAGAGAGATATCACCAACCAATCAACTTG 1249
DB 61 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 80
QY 1250 AGAAGAGGAGCCCGCTTCTTAACTTTGGAAGTTATTTGAGGTGAGGCCAGCAG 1309
DB 81 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 100
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGGAACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
DB 101 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 120
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGCTTATCGTGTCTGCTCAACAA 1429
DB 121 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 140
QY 1430 GGAACCTGGAAACCTTGAACCTCGTGGCTGTAAGAAAGAGCAACAACAGAGGGGACGGCGG 1489
DB 141 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 160
QY 1490 GAAGAAGAGGAGGAGCAAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1549
DB 161 GluGluGluGluAspGluAspGlnGluGluGlySerAsnArgGluValArgArgTyr 180
QY 1550 ACAGCGAGGTTGAAGAGGCGATGCTTCATCATGCGAGCAGCTCATCCAGTAGCCATC 1609
DB 181 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 200
QY 1610 AACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCAAGA 1669
DB 201 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 220
QY 1670 ATCTTCTTCAGGTGATAGGACAAATGTGATAGACAGATAGAGAGAGAGAGAGAGAGAT 1729
DB 221 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 240
QY 1730 TTAGCATTCCTCGGTGGGTGAACTTGAACCTTGAAGCTCATCAAAACCAAGAGAGGAACTCT 1789
DB 241 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnArgGluSer 260
QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCT 1849
DB 261 HisPheValSerAlaArgProGlnSerGlnSerProSer----- 273
QY 1850 TCTCCTGAGAAAGAGGATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1909
DB 274 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 293
QY 1910 ATTTTGAAGGCTTTTAACT 1927
DB 294 IleLeuLysAlaPheAsn 299
```

```
Q6PSU6
ID Q6PSU6 PRELIMINARY; PRT; 303 AA.
AC
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Conarachin (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang L., Liao B., Li H., Yan Y., Lin X., Huang S.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY581849; AAT00594.1; -.
DR GO; GO:0045735; F.nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 1.
FT NON TER 1
SQ SEQUENCE 303 AA; 34133 MW; 5B4E21BEB48654EC CRC64;

Alignment Scores:
Pred. No.: 7.84e-86 Length: 303
Score: 1435.00 Matches: 291
Percent Similarity: 95.50% Conservative: 6
Best Local Similarity: 93.57% Mismatches: 2
Query Match: 40.29% Indels: 12
DB: 2 Gaps: 4

US-10-728-051-1 (1-2032) x Q6PSU6 (1-303)
QY 1004 AATACGTGAGCGCGCTTCAATCGGAATTCATGAGATACGAGGCGTGTAGAA 1063
Db 2 AsnThrLeuGluAlaPheAsnAlaGluPheAsnGluLeuArgValLeuLeuGlu 21
QY 1064 GAGATCAGAGGTGAGCAGAGGAGAGAGGCGAGGCGATGAGTCTCGGAGTGT 1123
Db 22 GluAsnAlaGlyGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 41
QY 1124 GAGAACAAATGAGGAGGTAGTCAAGTGTCAAGAGGAGCAGTGTGAAGAACTTACTAAG 1183
Db 42 Asp---AsnGluGlyValIleValIleValIleValIleValIleValIleValIle 60
QY 1184 CAGCTAAATCGTCTCAAGAAAGGCTCCGAGAGAGGAGGATATCACCACCCCAATC 1243
Db 61 HisAlaLysSerValSerLysLysGlySerGluGluGluGluGluGluGluGluGlu 79
QY 1244 AACTTGAGAGAGCGGAGCGCGATCTTCTAACACTTTTCGGAGTATTATTGAGGTGAAG 1303
Db 80 AsnLeuArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLys 99
QY 1304 CCAGACAGAAGAACCCCGAGCTTCAGGACCTGACATGATGCTACCTGTGTAGAGATC 1363
Db 100 ProAspLysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIle 119
QY 1364 AAAGAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCTGCTC 1423
Db 120 LysGluGlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValVal 139
QY 1424 AACAAAGGAAGTGAACCTTGTAACCTGCTGCTGTGAAGAAAGCAACACAGAGGGGA 1483
Db 140 AsnLysGlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnAsgly 159
QY 1484 CGCGGGGAA-----GAAGAGGAGGACGAAGACGAAGAGAGGAGGGAAGTAACAGA 1534
Db 160 ArgArgGluGlnGluTrpGluGluGluGluGluGluGluGluGluGluGluGluGlu 179
QY 1535 GAGGTGCGTAGGTACACAGCGAGGTGAAGAGGCGATGTGTTTCATCATGCCAGCAGCT 1594

RESULT 7
Q7XXT2 PRELIMINARY; PRT; 621 AA.
AC
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Prepro beta-conglycinin alpha prime subunit.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Urade R., Nakatani H., Nakano C.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB113351; BAC78524.1; -.
DR PDB; 1UIK; X-ray; A/B/C=--
DR GO; GO:0045735; F.nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 2.
SQ SEQUENCE 621 AA; 72247 MW; 0DF4B6E337337B7F CRC64;

Alignment Scores:
Pred. No.: 5.1e-77 Length: 621
Score: 1302.00 Matches: 286
Percent Similarity: 61.02% Conservative: 121
Best Local Similarity: 42.88% Mismatches: 170
Query Match: 36.55% Indels: 90
DB: 2 Gaps: 19

US-10-728-051-1 (1-2032) x Q7XXT2 (1-621)
QY 50 ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValValPheLeuAlaSerVal 19
QY 110 TTGCAACCGATGCCAAGTCATCCTTACCAAGAAACACAGAACCCCTCGGCCAG 169
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GlnAsnProSerHisAsn 36
QY 170 AGGTGCTCCAGATGTTGTCACAGGAAACCGGAGTCTTACGAGCAAGGAGGAGGAGTCT 229
Db 170 AGGTGCTCCAGATGTTGTCACAGGAAACCGGAGTCTTACGAGCAAGGAGGAGGAGTCT 229
```

Db 37 LysCysLeuArgSerCysAsnSerGluLysAspSerTyrArgAsnGlnAlaCysHisAla 56  
QY 230 CGCTGCACC-----AAGCTCGAGTATGATCCTCGTTGT-----GTCATGAT 271  
Db 57 ArgCysAsnLeuLeuLysValGluGluGluGluGluCysGluGluGlnIleProArg 76  
QY 272 CTTCTGA---GGACACACTGCGCACCAACCAACCACTTCCCTCCAGGGGAGGGACACGT 328  
Db 77 ProArgProGlnHisProGluArgGluArgGlnGlnHisGlyGluLysGluGluAspGlu 96  
QY 329 GGCCGCCAACCCGGAGACTACGATGATGACCGC---CGTCAACCCCGAAGAGAGAA--- 382  
Db 97 GlyGluGlnProArgProPheProPheProArgGlnProArgGlnGluGlu 116  
QY 383 -----GGAGSCCGATGG 394  
Db 117 HisGluGlnLysGluGluHisGluTrpHisArgLysGluGluLysHisGlyLysGly 136  
QY 395 GGA-----CCAGCTGGACCGAGGAGCGGTGAA 421  
Db 137 SerGluGluGlnAspGluArgGluHisProArgProHisGlnProHisGlnLysGlu 156  
QY 422 AGAGAGAA---GACTGGAGAACCAAGAGAGATTTGGAGCGCAACAGTCAATCAGCAG 478  
Db 157 GluGluLysHisGluTrpGlnHisLysGlnGluLysHisGlnGlyLysGluSerGluGlu 176  
QY 479 CCACGGAATAAGGCCGAGGAGGAGAGAGAGAGAAAGAG---TGGGGAACACCCAGGT 535  
Db 177 GluGluLysGlnAspGluAspGluGluGlnAspLysGluSerGlnGluSerGluGly 196  
QY 536 AGCCATGTGAGGAGAAACATCTCGAAC-----AACCTTTCTACTTCCGTCA 586  
Db 197 SerGluSerGlnArgGluProArgArgHisLysAsnLysAsnProPheHisPheAsnSer 216  
QY 587 AGCGGTTAGCACCCCTAGCGGAGCAACCAACCGTAGGATCGGCTCTGCGAGAGTTT 646  
Db 217 LysArgPheGlnThrLeuPheLysAsnGlnTyrGlyHisValArgValLeuGlnArgPhe 236  
QY 647 GACCAAGGTCAAGGCAAGTTTCAGAAATCCAGAAATCAGGATTCAGATCGAGGCG 706  
Db 237 AsnLysArgSerGlnGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSer 256  
QY 707 AACCTTAACACTCTTGTCTTCCAGACGCTGATGCTGATTAACATCTTGTATCCAG 766  
Db 257 LysProAsnThrLeuLeuLeuProHisHisAlaAspAlaAspTyrLeuIleValIleLeu 276  
QY 767 CAAGGCAACCCCGTGCCTGAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGAC 826  
Db 277 AsnGlyThrAlaIleLeuThrLeuValAsnAsnAspAspArgAspSerTyrAsnLeuGln 296  
QY 827 GAGGCGCATGCACTCAGAAATCCCGGTTTCATTTCTTACATCTTGAACCGCCATGAC 886  
Db 297 SerGlyAspAlaLeuArgValProAlaGlyThrThrTyrTyrValValAsnProAspAsn 316  
QY 887 AACCAAGAACTCAGTAGTAGTAAATCTCCATGCTCCGCTTAACACACCCGCGCAGTTTGAG 946  
Db 317 AspGluAsnLeuArgMetIleThrLeuAlaIleProValAsnLysProGlyArgPheGlu 336  
QY 947 GATTTCTTCCCGGAGCAGCGAGCAATCATCTTCTTCTTCTTCTTCTTCTTCTTCTT 1006  
Db 337 SerPhePheLeuSerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerLysAsn 356  
QY 1007 ACCTTGGAGGCGCTTCAATCGGAAATCAATGAGATACGGAGGGTCTGTTAGAGAG 1066  
Db 357 IleLeuGluAlaSerTyrAspThrLysPheGluGluIleAsnLysValLeuPheGlyArg 376  
QY 1067 AATCAGAGGTGAGCAA---GAGGAGAGGGCAGGCGATGCGGATCTCGGAGTAGT 1123  
Db 377 GluGluGlyGlnGlnGlnGlyGluGluArgLeuGln----- 388  
QY 1124 GAGAACAAATGAGGAGGTAGTAGTCAAGAGTCAAGAGGACGAGTGAAGAACTTACTTAAG 1183  
Db 389 -----GluSerValIleValGluIleSerLysGlnIleArgGluLeuSerLys 405

QY 1184 CACGCTAAATCCGTCTCAAAGAAAGGCTCCGAAGAGAGGAGATATACCAACCAATC 1243  
Db 406 HisAlaLysSerSerSerArgLysThrIleSerSerGlu-----AspLysProPhe 422  
QY 1244 AACTTGGAGAGAGGAGCGCGATCTTTCTTAACAACTTTGGAGAGTTATTTGAGTGAG 1303  
Db 423 AsnLeuArgSerArgAspProIleTyrSerAsnLysLeuGlyLysLeuPheGluIleThr 442  
QY 1304 CCAGACAGAGAGAACCCCGAGCTTCAGACCTTCAGACATGATCTCACCTGTGTAGAGATC 1363  
Db 443 ProGlu---LysAsnProGlnLeuArgAspLeuAspValPheLeuSerValValAspMet 461  
QY 1364 AAAGAGAGAGCTTTGATGCTCCCACTTCAACTCAAGGCCATGTTATCGTCTGCTC 1423  
Db 462 AsnGluGlyAlaLeuPheLeuProHisPheAsnSerLysAlaIleValValLeuValIle 481  
QY 1424 AACAAAGGAACTGGAAACCTTGAACCTGCTGCTGTGTAAGAAAGACCAACAGAGGGA 1483  
Db 482 AsnGluGlyGluAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnArgGln 500  
QY 1484 CGGCGGAGAGAGAGGAGGAGCAAGACGAAGAGAGGAGGAAGTAACAGAGAGGTGGCT 1543  
Db 501 GlnGlnGluGluGlnPro-----LeuGluValArg 510  
QY 1544 AGGTACACAGCGAGGTTGAAGAGCGGATGTTTCATCATGCTCCAGAGCTCATCCAGTA 1603  
Db 511 LysTyrArgAlaGluLeuSerGluGlnAspIlePheValIleProAlaGlyTyrProVal 530  
QY 1604 GCATCAACGCTTCTCCGAACTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1663  
Db 531 ValValAsnAlaThrSerAspLeuAsnPhePheAlaPheGlyIleAsnAlaGluAsnAsn 550  
QY 1664 CACAGAACTCTCTTGCAGGTGATTAAGGACAAATGTGATAGACAGATAGAGCAAGCG 1723  
Db 551 GlnArgAsnPheLeuAlaGlySerLysAspAsnValIleSerGlnIleProSerGlnVal 570  
QY 1724 AAGGATTTAGCAATCTCCTGGTGGGTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1783  
Db 571 GlnGluLeuAlaPheProGlySerAlaLysAspIleGluAsnLeuIleLysSerGlnSer 590  
QY 1784 GAATCTCACTTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1843  
Db 591 GluSerTyrPheValAspAlaGlnProGln----- 600  
QY 1844 AAAGAGTCTCTCGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGAGGAG 1903  
Db 601 -----GlnLysGluGluGlyAsnLysGlyArgLysGlyProLeu 613  
QY 1904 CTTTCAATTTGAAGCTTTT 1924  
Db 614 SerSerIleLeuArgAlaPhe 620  
RESULT 8  
Q6EBCL PRELIMINARY; PRT; 533 AA.  
ID Q6EBCL; AC Q6EBCL; DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Beta-conglutinin.  
OS Lupinus albus (white lupine).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.  
OX NCBI\_TaxID=3870;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Monteiro S.A., Freitas R.M., Teixeira A.N., Ferreira R.B.;  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY500372; AAS97865.1; -  
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
DR InterPro; IPR006045; Cupin.





DE Beta-conglycinin alpha prime subunit.  
OS Glycine max (Soybean)  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fukazawa C.;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB030838; BAB64303.1; -;  
DR HSSP; P25974; 11PJ.  
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
DR InterPro; IPR006045; Cupin.  
DR InterPro; IPR007113; Cupin region.  
DR InterPro; IPR01051; RmlC\_Like\_Cupin.  
DR Pfam; PF00190; Cupin; 2.  
SQ SEQUENCE 621 AA; 72138 MW; 0196CAD3C6B566DC CRC64;

Alignment Scores:  
Pred. No.: 7,76e-76 Length: 621  
Score: 1284.00 Matches: 284  
Percent Similarity: 59.56% Conservative: 118  
Best Local Similarity: 42.07% Mismatches: 167  
Query Match: 36.05% Indels: 106  
DB: 2 Gaps: 19

US-10-728-051-1 (1-2032) x Q948Y0 (1-621)

QY 50 ATGAGAGGGAGGTTTCTCCAGTGTCTTCTAGGATCCTGTCTGGTTCAGTT 109  
DB 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValPheLeuAlaSerVal 19

QY 110 TCTGCAAGCATGCCAGTATCATCTTACAGAGAAACAGAACCCCTCGGCCAG 169  
DB 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GlnAsnProSerHisAsn 36

QY 170 AGGTGCTCCAGAGTCTTCAACAGGACCGATGATCTTGAAGCAAAAGGATCGAGTCT 229  
DB 37 LysCysLeuArgSerCysAsnSerGluLysAspSerTyrArgAsnGlnAlaCysHisAla 56

QY 230 CGTGTGACC-----AGCTCGATGATGATCCTGTTGT-----GTCATGAT 271  
DB 57 ArgCysAsnLeuLeuLysValGluGluGluCysGluGluGluGlnIleProArg 76

QY 272 CCTCGA---GGACACTGGCACCACCAACCAACCTCCCTCCAGGGAGCGACACGT 328  
DB 77 ProArgProGlnHisProGluArgGluArgGlnHisGlyGluLysGluLysGlu 96

QY 329 GGCCGCCAACCCGGAGACTACGATGATGACCGC---CGTCAACCCCGAAGAGAGGA 385  
DB 97 GlyGluGlnProArgProPheProArgProArgProArgGlnProArgGlnGlu----- 114

QY 386 GGCCGATGGGACACAGCTGGACCGAGGAGCGTGAAGAGAGAGACTGGACACCA 445  
DB 115 -----GlyGluHisGluLysGluLysGluLysGluLysGluLys 128

QY 446 AGAGAA-----GATTGGAGGGGACCAAGTATCATGACGACCCAGG 484  
DB 129 GluGluLysHisGlyGlyLysGlySerGluGluGluGlnAspGlyArgGluHisProArg 148

QY 485 AAAATAAGGCC----- 496  
DB 149 ProHisGlnProHisGlnLysGluGluGluLysHisGluTrpGlnHisGlnGluLys 168

QY 497 ---GAAGAGAGAGAGGACACAG----- 520  
DB 169 HisGlnGlyLysGluSerGluGluGluGluGluAspGlnAspGluAspGluGlnAsp 188

QY 521 -----TGGGGAACACAGGATGACCATGTGAGGGAAGAAACATCTCGAAC----- 565  
DB 189 LysGluSerGlnGluSerGluGlySerGluSerGluSerGluArgGluHisLysAsn 208

QY 566 ---AACCCCTTTCTACTCCCGTCAAGCGGCTTTAGCACCCCGCTACGGGAACCAAAACCGT 622  
DB 209 LysAsnProPheHisPheAsnSerLysArgPheGlnThrLeuPheLysAsnGlnTyrGly 228

QY 623 AGGATCCGGTCTCCAGAGGTTTTCACCAAGAGTCAAGGCGAGTTTCAGAAATCCAGAA 682  
DB 229 HisValArgValLeuGlnArgPheAsnLysArgSerGlnGlnLeuGlnAsnLeuArgAsp 248

QY 683 CACCGTATTGTGCAGATCGAGCCCAAACTTAACACTCTTGTCTTCTCCCAAGCACCTCAT 742  
DB 249 TyrArgGlnLeuGluPheAsnSerLysProAsnThrLeuLeuLeuProHisHisAlaAsp 268

QY 743 GCTGATAACATCTTGTATTATCCAGACAGGCAAGCACCGTACGTCAGTACAAATCGCAAT 802  
DB 269 AlaAspTyrLeuIleValIleLeuAsnGlyThrAlaIleLeuThrLeuValAsnAsnAsp 288

QY 803 AACAGAAAGAGCTTTAATCTTCCAGAGGCCATGCATCCAGATCCCATCCGGTTTCATT 862  
DB 289 AspArgAspSerTyrAsnLeuGlnSerGlyAspAlaLeuArgValProAlaGlyThrThr 308

QY 863 TCCTATCATCTTGAACCGCATGACAAACAGAACCTTCAGATGATGATTAATCTCCATGCC 922  
DB 309 TyrTyrValValAsnProAspAsnAspGluAsnLeuArgMetIleThrLeuAlaIlePro 328

QY 923 GTTAACACACCCGGCGCATTTGAGGATTTCTCCCGCGAGCAGCGGAGACCAATCATCC 982  
DB 329 ValAsnLysProGlyArgPheGluSerPhePheLeuSerSerThrGlnAlaGlnSer 348

QY 983 TACTTGCAGGCTTCAGCAGGAATACGTTGGAGGCGCTTCAATCGCGAATTCATAG 1042  
DB 349 TyrLeuGlnGlyPheSerLysAsnIleLeuGluAlaSerTyrAspThrLysPheGluGlu 368

QY 1043 ATACGAGGCTCTGTTAGAGAGAAATGAGGAGGTGAGCAA---GAGGAGAGAGGCGAG 1099  
DB 369 IleAsnLysValLeuPheGlyArgGluGluGlyGlnGlnGlnGlyGluArgLeuGln 388

QY 1100 AGCGATGAGTACTCGGAGTAGTGAGACAATGAAGAGTGATGATGATGATGATGATG 1159  
DB 389 -----GluSerValIleValGluLysSerLys 397

QY 1160 GAGCAGTTGAAGAACTTACTAAGCAGCTAAATCCGTCTCAAGAAAGCGTCCGAAGAA 1219  
DB 398 LysGlnIleArgGluLeuSerLysArgAlaLysSerSerSerArgLysThrIleSerSer 417

QY 1220 GAGGAGATATCACAACCAATCACTTGAGAGAGGCGGAGCCGATCTTTCTTAACAAC 1279  
DB 418 Glu-----AspLysProPheAsnLeuArgSerArgAspProIleTyrSerAsnLys 434

QY 1280 TTTGGGAAGTTATTGAGTGAAGCGCAGACAGAGAACCCCGCTTCAGGACCTGGAC 1339  
DB 435 LeuGlyLysLeuPheGluIleThrProGlu---LysAsnProGlnLeuArgAspLeuAsp 453

QY 1340 ATGATCTCACCTGTGTAGAGATCAAGAGAGAGCTTTGATGCTCCACACTTCAACTCA 1399  
DB 454 ValPheLeuSerValValAspMetAsnGluGlyAlaLeuPheLeuProHisPheAsnSer 473

QY 1400 AAGCCATCGTTATCGTCTGTCACAAAGAGAACTGGAAACCTTGAACCTCGTGGCTGTA 1459  
DB 474 LysAlaIleValValLeuValIleAsnGluGlyGluAlaAsnIleGluLeuValGlyIle 493

QY 1460 AGAAAGACCAACACAGAGGGGCGCGGAGAGAGAGAGGAGGAGCAAGCAAGAAAGAG 1519  
DB 494 ---LysGluGlnGlnArgGlnGlnGlnGluGluGlnPro----- 506

QY 1520 GAGGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTGTAAGGAAGCGATGTTC 1579  
DB 507 -----LeuGluValArgLysTyrArgAlaGluLeuSerGluGlnAspIlePhe 522

QY 1580 ATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTCCGAACCTCATCTGCTTGGC 1639  
DB 523 ValIleProAlaGlyTyrProValValValAsnAlaThrSerAspLeuAsnPhePheAla 542

QY 1640 TTCGGTATCAACGCTGAAACACACCAACAGATCTCTCTTCAGGATGATAAGGACAATGTG 1699

```
Db 543 PheGlyIleAsnAlaGluAenAenGlnArgAsnPheLeuAlaGlySerLysAspAenVal 562
Qy 1700 ATAGACAGATAGAGAGCAAGCAAGGATTAGCATTCCTCGGTGGGTGAACAAGTT 1759
Db 563 IleSerGlnIleProSerGlnValGlnGluLeuAlaPheLeuGlySerAlaLysAspIle 582
Qy 1760 GAGAGCTCATCAAAAACCAAGAGGAATCTCATTGTGTAGTCTGCTCTCAATCTCAA 1819
Db 583 GluAsnLeuIleLysSerGlnSerGluSerTyrPheValAspAlaGlnProGln----- 600
Qy 1820 TCTCAATCTCCGTCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGATCAAGAGGAGGAA 1879
Db 601 -----GlnLysGluGluGly 605
Qy 1880 AACCAAGAGGAGGAGGTCCACTCTTCAATTTGAAGGCTTTT 1924
Db 606 AsnLysGlyArgLysGlyProLeuSerSerIleLeuArgAlaPhe 620

RESULT 10
Q94LX2 PRELIMINARY; PRT; 605 AA.
AC Q94LX2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-conglycinin alpha subunit.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids;
OC Eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21327318; PubMed=11434464;
RA Yoshino M., Kanazawa A., Tsutsuni K., Nakamura I., Shimamoto Y.;
RT "Structure and characterization of the gene encoding alpha subunit of
RT soybean beta-conglycinin."
RL Genes Genet. Syst. 76:99-105(2001).
DR EMBL; AB051865; BAB56161.1; -.
DR PIR; S20007; S20007.
DR HSSP; P25974; 1IPJ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 2.
SQ SEQUENCE 605 AA; 70306 MW; 8ACE6F8532662984 CRC64;

Alignment Scores:
Pred. No.: 1.04e-75 Length: 605
Score: 1282.00 Matches: 285
Percent Similarity: 59.37% Conservative: 108
Best Local Similarity: 43.05% Mismatches: 173
Query Match: 35.99% Indels: 96
DB: 2 Gaps: 18

US-10-728-051-1 (1-2032) x Q94LX2 (1-605)
Qy 50 ATGAGAGGAGGGTTTCTCCACTGATCTGCTAGGATCCTTGCTCGGCTTCAGTT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
Qy 110 TCTGCAACGATGCCAAGTCATCACCTTACCAGAGAAACACAGAACCCCTCGGCCAG 169
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36
Qy 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGAGTCT 229
Db 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
Qy 230 CGTGGCAACCAAGCTCGAGTATGAT-----CCTCGTTGT 262
|||||
|||||
```

```
Db 57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyGluIleProArg--- 75
Qy 263 GTCTATGATCTTCGAGACACACTGGCACCAACCAACGTTCCCTTCAGGAGCGG 322
Db 76 -----ProArgProArg-----ProGlnHisProGlu 84
Qy 323 ACAGTGGCGCCCAACCCGAGACTACGATGATGAC----- 358
Db 85 ArgGluProGlnGlnProGlyGluLysGluGluAspGluGlnProArgProIle 104.
Qy 359 -----CGCGCTCAACCCCGAAGAGAG-----GAAGGAGGCCGA 391
Db 105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGluGlnGlu 124
Qy 392 TGG-----GCACGACTGGACCCGAGGAGCGGTGAAGAGAGAA 430
Db 125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGlu 144
Qy 431 GACTGGAGACNACCAAGAGAGAT-----TGAGGGGACCAAGTCTATCAGCAGCCACGG 484
Db 145 AspGluGluGlnAspGluArgGlnPheProPheProArgProHisGlnLysGluGlu 164
Qy 485 AAAATAAGGCCCGCAAGAGAGAGAGAGAGAGAGTGGGGAAACACCAGGTAGCCATGTG 544
Db 165 ArgLysGlnGluGluAspGluAspGluGlnGlnArgGluSerGluGluSerGluAsp 184
Qy 545 AGGGAA-----GAAACATCTCGGAACAACCTTTCTACTTCCGTCACAGCGGTTTACG 598
Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
Qy 599 ACCCGTACGGGAACCAAAACGGTAGGATCCGGGCTCTGCAGAGGTTTGCACCAAGGTCA 658
Db 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
Qy 659 AGCGAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTTAACACT 718
Db 225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleGluGluPheAsnSerLysProAsnThr 244
Qy 719 CTTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATCCAGCAAGGCGCAAGCC 778
Db 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
Qy 779 ACCGTACCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCA 838
Db 265 IleLeuSerLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
Qy 839 CTCAGAAATCCCATCCGCTTTCATTTCTTACATCTTGAACCGCATGACCAACCAAGACCTC 898
Db 285 LeuArgValProSerGlyThrThrTyrValValAsnProAsnAsnGluAsnLeu 304
Qy 899 AGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCGCGCAGTTTGAGGATTTCTCCCG 958
Db 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 324
Qy 959 GCGAGACGCGGAGACCAATCATCTTCTGAGGGCTTTCAGAGGAATACGTTGGAGGCC 1018
Db 325 SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
Qy 1019 GCCTTCAATCGCGAATTCATGAGATACGAGGCTGTCTTTAGAGAAGATGCGAGGAGGT 1078
Db 345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly--- 363
Qy 1079 GAGCAAGAGAGAGAGGAGCGGATGAGTACTCGGAGTAGTGAGAAACAATGAAGGA 1138
Db 364 -----GlnGlnGlnGlyGlnArgLeu-----GlnGluSer 374
Qy 1139 GTGATAGTCAAGGTCAAAAGGAGCAGCTTGAAGAAGACTTACTAAGCAGCGTAATCCGTC 1198
Db 375 ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
Qy 1199 TCAAGAAAGGCTCCGAGAGAGGAGGAGATATCACCAACCAATCACTTGCAGAGAGGC 1258
Db 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
|||||
|||||
```

```
QY 1259 GAGCCCGATCTTTCTAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAGAGAAGAAC 1318
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 412 AppProLety:SerAsnLysLeuGlyLysPhePheGluLeuThrProGlu---LysAsn 430
QY 1319 CCCAGCTTCAGACCTGACATGATGTCACCTGTGTAGATCAAGAGAGAGCTTTG 1378
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 431 ProGlnLeuArgAspLeuAspLeuPheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCACACTCAACTCAAGGCGCATGTTATCGTCTGCTCAACAAAGGAACGGA 1438
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTGAACCTCGCTGTGTAAGAAAGAGACACAGAGGCGCGGGAAGAGAG 1498
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGluGln 489
QY 1499 GAGGCGAAGACGAAAGAGAGGAGGAAGTAACAGAGAGGTGGTACACAGCGAGG 1558
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
QY 1559 TTGAAGGAGCGGATGTTTCATCGCCAGAGCTCATCCAGTACGCCATCAACGCTCC 1618
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
QY 1619 TCGGAATCTCATCTGCTTGGCTTCGCTATCAACGCTGAAACACACAGAAATCTTCTT 1678
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnGlnArgAsnPheLeu 539
QY 1679 GCAGGTGATAAGACATGTGTATGACACAGATAGAGAACGCGAAGATTTAGCATTC 1738
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
QY 1739 CTGGGTGGGTCAACAGTTGAGAGCTCATCAAAACACAGAGGAATCTCACTTTGG 1798
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
QY 1799 AGTGCTGCTCTCAATCTCAATCTCAATCTCGCTGCTCTGAGAAAGAGTCTCTGAG 1858
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 580 AspAlaGlnProLysLys----- 585
QY 1859 AAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1918
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
QY 1919 GCTTTT 1924
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 603 AlaPhe 604

RESULT 11
GLCA_SOYBN
ID GLCA_SOYBN STANDARD; PRT; 605 AA.
AC P13916;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta-conglycinin, alpha chain precursor.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Cotyledon;
RX MEDLINE=9135860; PubMed=2103438;
RA Sebastiani F.L., Farrel L.B., Schuler M.A., Beachy R.N.;
RT "Complete sequence of a cDNA of alpha subunit of soybean beta-
RT conglycinin."
RL Plant Mol. Biol. 15:197-201(1990).
CC -|- FUNCTION: Seed storage protein. Accumulates during seed
CC development and is hydrolyzed after germination to provide a
CC carbon and nitrogen source for the developing seedling.
```

```
CC -|- SUBUNIT: The alpha-, alpha-, and beta-subunits associate in
CC various combinations to form trimeric proteins.
CC -|- SUBCELLULAR LOCATION: Embryo axis, and cotyledonary membrane-bound
CC vacuolar protein bodies.
CC -|- SIMILARITY: Belongs to the 7S seed storage protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17698; CAA35691.1; -.
DR PIR; S14681; FWSVBA.
DR HSSP; P25974; IIRJ.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011031; Multihaem cyt.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 2.
KW Glycoprotein; Multigene family; Seed storage protein; Signal.
FT SIGNAL 1..22
FT PROPEP 23..62
FT CHAIN 63..605 Beta-conglycinin, alpha chain.
FT CARBOHYD 261..261 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 517..517 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 605 AA; 70293 MW; CBEB30506BBBC57 CRC64;

Alignment Scores:
Pred. No.: 1,218-75 Length: 605
Score: 1281.00 Matches: 285
Percent Similarity: 59.21% Conservative: 107
Best Local Similarity: 43.03% Mismatches: 174
Query Match: 35.96% Indels: 96
DB: 1 Gaps: 18

US-10-728-051-1 (1-2032) x GLCA_SOYBN (1-605)
QY 50 ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTGCTAGGATCCTTGTCTGCTTCAGTT 109
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
QY 110 TCTGCAACGATGCCAAGTCACTACCTTACAGAGAAGAAACAGAGAACCCCTGCGCCAG 169
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36
QY 170 AGTGCTCTCAGATGTTGTCAACAGGACCGGATGACTTGAAGCAAAAGCATCGAGTCT 229
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGCTGCACCAAGCTCGAGTATGAT-----CCTCGTTGT 262
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyGluLeuProArg--- 75
QY 263 GTCTATGATCTCTGAGGACACACTGSCACCAACCAACCAACCTTCCCTCCAGGGAGCG 322
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 76 -----ProArgProArg-----ProGlnHisProGlu 84
QY 323 ACAGTGGCGGCGCAACCGGAGACTACGATGATGAC----- 358
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 85 ArgGluProGlnGlnProGlyGluLysGluGluAspGluAspGluGlnProArgProIle 104
QY 359 -----CGCGCTCAACCCCGAAGAGAG-----GAAGGAGGCCGA 391
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 105 ProPheProArgProGlnProArgGlnGluGluGluHisGluGlnArgGluGlnGlu 124
QY 392 TGG-----GGACCGCTGGACCGGAGGCGTGAAGAGAGAA 430
DB      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 125 TrpProArgLysGluGluLysArgGlyGlyLysGlySerGluGluGluAspGlu 144
QY 431 GACTGAGACACCAAGAGAGAGAT-----TGGAGCGACCAAGTATCATCAGCAGCCACGG 484
```

Dbb

LeuGluValArgLysTyrArgAlaGlu 499

TTCAAGGAAGGGCATGTGTTTCATCGCCACGACTCCTCACGTAGGCATCAACGCTTCC 1618  
||||| ||||| :::::||||| :||| |::| |::| |::| |::| |::|  
500 LeuSer-GluGlnAspilePheValIleProAlaGlyTy-ProValValValAsnAlaThr 519

TCCGAATCTCATCTGCTTGCTTCGGTATCAAACGCTGAAACAACCACACAGAATTCTTCCTT 1678  
||||| ::| :::  
520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnGlnArgAsnPheLeu 539

GCGAGTGATAAAGCACCAATGTGATAGAACCATAGAGAAGCAACGGAAGGATTTTAGCATTC 1738  
||||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|  
540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559

CCTGGGTCGGGTGAACAAGTTGAGAAGCTCATCAAAAAACAGAGAAGTAATCTCACTTTGTG 1798  
||||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|  
560 ProGly-Ser-AlaGlnAlaValGluLyLeuLeuLyAsnGlnArgGluSertYrPheVal 579

AGTCTCGTCTCCAATCTCAATCTCCTGCTGCTCTCCTGAGAAGAAGTCTCTCTGAG 1858  
||||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|  
580 AspAlaGlnProLyslys- - - - - 585

AAAGAGGATCAAGACGAGGAGAAAACCAAGGAGGAGGTCCACTCTCTTCAATTTTGAAG 1918  
::::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|  
586 -----LysGluGluGlyAsnLyAsnGlyArgylscglyProLeuserrilleLeuarig 602

GCTTTT 1924  
|||||  
603 AlaPhe 604

RESULT 12

CVCA\_PEA STANDARD; PRT; 571 AA.

ID CVCA\_PEA ID AC PI3915;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 03-JUL-2004 (Rel. 44, Last annotation update)  
DE Convicilin precursor.  
GN Name=CVCA;  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids;  
OC eurosid1; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
NCBI\_TaxID=388;  
RN NCBI TaxID=388;  
[1]  
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=cv\_Felham First;  
RX MEDLINE=88326208; PubMed=3415641;  
RA Bown D., Ellis T.H.N., Gatehouse J.A.;  
RT "The sequence of a gene encoding convicillin from pea (*Pisum sativum*) L.) shows that convicilin differs from vicilinin by an insertion near the N-terminus";  
RL Biochem. J. 251:717-726(1988).  
CC CC  
CC !- FUNCTION: Seed storage protein.  
CC !- SUBCELLULAR LOCATION: Cotyledonary membrane-bound vacuolar protein bodies.  
CC  
CC !- SIMILARITY: Belongs to the 7S seed storage protein family.  
  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.ebi.ac.uk/termsandconditions) or send an email to licensee@ebi.ac.uk).

EMBL; X06398; CAA28695.1; --  
DR PIR; S00566; S00566.  
DR HSPR; P25974; IIPJ.  
DR InterPro; IPRO06045; Cupin.  
DR InterPro; IPRO07113; Cupin region.  
DR InterPro; IPRO11051; Rmlc\_Like\_cupin.  
DR Pfam; PF00190; Cupin; 2.

KW Direct protein sequencing; Multigene family; Seed storage protein;

KW Signal. 1 28 Convicillin.  
FT CHAIN 29 571  
SQ SEQUENCE 571 AA; 66989 MW; 749CFBEB2D16D57B CRC64;

## Alignment Scores:

Pred. No.: 9.95e-75 Length: 571  
Score: 1267.00 Matches: 276  
Percent Similarity: 56.67% Conservative: 98  
Best Local Similarity: 41.82% Mismatches: 146  
Query Match: 35.57% Indels: 140  
DB: 1 Gaps: 13

US-10-728-051-1 (1-2032) x CVCA\_PBA (1-571)

QY 68 CCACTGATGCTGTGTCAGGATCCCTGCTGCTTCTGCAACGCATGCCAAG 127  
DB 10 ProLeuLeuLeuPheLeuGlyIlePheLeuAlaSerValCysValThrTyrAlaAsn 29  
QY 128 TCATCCTTTACCAGAGAAAAACAGAGAACCCCTGCGCCAGAGGTGCTCCAGAGTTGT 187  
DB 29 ----- 29  
QY 188 CAACAGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGTGCACCAAGCTCGAG 247  
DB 29 ----- 29  
QY 248 TATGATCCTGTTGTGCTATGATCCTCGAGGACACTCGGACACCAACCAACGTTCC 307  
DB 30 -----tyrAspGluGlySerGluThrArgVal----- 38  
QY 308 CCTCCAGGAGGGGACACGCTGGCCGCAACCCGAGACTPACCATGATGACCGCCGTCAA 367  
DB 39 ---ProGlyGlnArgGluArgGlyArgGlnGluGlyGluGlyGluGlyArgHis--- 56  
QY 368 CCCCAGAGAGGAGGAGCGGATGGGACCGAGTGGACCGAGGAGCGT----- 418  
DB 57 -----GlyGluTrpArgProSerTyrGluLysGluGluHisGluGlu 70  
QY 418 ----- 418  
DB 71 GluLysGlnLysTyrArgTyrGlnArgGluLysGluLysGluValGlnProGly 90  
QY 419 -----GAAAGAGAGAGAGACTCGAGACAACCAAGAGAGATTGGAG----- 460  
DB 91 ArgGluArgTrpGluArgGluGluAspGluGluGlnValGluGluTrpArgGlySer 110  
QY 461 -----CGACCAAGTCTATCAGCAGCCACGG 484  
DB 111 GlnArgArgGluAspProGluGluArgAlaArgLeuArgHisArgGluGluArgThrLys 130  
QY 485 AAAATAGGCCCGAAGGAGAGAGAGAGAGAGAGTGGGGAACACCGAGTACCATGTG 544  
DB 131 ArgAspArgArgHisGlnArgGluGluGlyGluGlu-----GluArgSer 145  
QY 545 AGGAGAGAAACATCTCGGAACAAACCTTCTACTTCCCGTCAAGGGGTTTAGCACCCGC 604  
DB 146 SerGluSerGlnGluHisArgAsnProPheLeuPheLysSerAsnLysPheLeuThrLeu 165  
QY 605 TAGGGNACCAAAACGGTAGGATCCGGGTCCGAGAGTTTGACCAAGGTCAAGGCAG 664  
DB 166 PheGluAsnGluAsnGlyHisIleArgArgLeuGlnArgPheAspLysArgSerAspLeu 185  
QY 665 TTTCAGATCTCCAGATCACCTATTGTGTCAGATCGAGGCCAAACCTAACACTCTTGT 724  
DB 186 PheGluAsnLeuGlnAsnTyrArgLeuValGluTyrArgAlaLysProHisThrIlePhe 205  
QY 725 CTTCCCAAGCACGCTGATGCTGATAACATCTTTGTTATCCAGCAAGGCGCAACCGGTG 784  
DB 206 LeuProGlnHisIleAspAlaAspLeuIleLeuValValLeuAsnGlyLysAlaIleLeu 225  
QY 785 ACCGTAGCAATGCCAATAACAGAGAGCTTTAATCTTGACGAGGGCCATGCACCTCAGA 844

DB 226 ThrValLeuSerProAsnAspArgAsnSerTyrAsnLeuGluArgGlyAspThrIleLys 245  
QY 845 ATCCCATCGGTTTCTTCTTACATCTTGAACCGCATCAACACAGAACCTTCACAGTA 904  
DB 246 IleProAlaGlyThrThrSerTyrLeuValAsnGlnAspGluGluAlaSerVal 265  
QY 905 GCTAAATCTCCATGCCCTTAAACACACCGCCGACGTTTGTAGAGATTCTTCGCCGCGAGC 964  
DB 266 ValAspPheValIleProValAsnArgProGlyLysPheGluAlaPhe-----GlyLeu 283  
QY 965 AGCCGAGACCAATCATCTTCTTGAAGGCTTCAGCAGGAATACGTGTGAGCGCGCTTC 1024  
DB 284 SerGluAsnLysAsnGlnTyrLeuArgGlyPheSerLysAsnIleLeuGluAlaSerLeu 303  
QY 1025 AATGCCGAATTCATGAGATACGGAGGTGCTGTGTAGAGAGATTCAGGAGGTGAGCA 1084  
DB 304 AsnThrLysTyrGluThrIleGluLysValLeuLeuGluGlnGluLysProGln 323  
QY 1085 GAGGAGAGGGGACAGGCGATCGAGTACTCGAGTAGTGAGAACAAATGAAGGAGTGATA 1144  
DB 324 GlnLeuArgAspArgLysArg-----ThrGlnGlnGlyGluGluArgAsp---AlaIle 340  
QY 1145 GTCAAAGTCTCAAAGAGACGTTTGAAGAACTTACTAAGCACGCTTAATCTCTCAAG 1204  
DB 341 IleLysValSerArgGluGlnIleGluGluLeuArgLysLeuAlaLysSerSerLys 360  
QY 1205 AAAGGCTCCGAGAGAGGGAGATATCAACCAACCCCACTTGAAGAGAGCGGAGCCC 1264  
DB 361 LysSerLeuProSerGluPheGlu-----ProPheAsnLeuArgSerHisLysPro 377  
QY 1265 GATCTTTTAAACAACCTTTGGGAAGTTATTGTAGGTGAAGCCAGACAAGAGAACCCCCAG 1324  
DB 378 GluTyrSerAsnLysPheGlyLysLeuPheGluIleThrProGluLysLysTyrProGln 397  
QY 1325 CTTGAGGACCTCGACATGATGCTCACCTGTGTAGAGATCAAAGAGAGAGCTTTGATGCTC 1384  
DB 398 LeuGlnAspLeuAspIleLeuValSerCysValGluIleAsnLysGlyAlaLeuMetLeu 417  
QY 1385 CCACACTTCAACTCAAGGCCCATGTTTCTGCTCGTCAACCAAGGAACCTGAAACCTT 1444  
DB 418 ProHisTyrAsnSerArgAlaIleValLeuLeuValAsnGluGlyLysGlyAsnLeu 437  
QY 1445 GAACCTGTGCTGTAAAGAAAGACCAACACAGAGGGGACGCGGGAAGAGAGGAGAC 1504  
DB 438 GluLeuLeuGlyLeuLysAsnGluGlnGluArg----- 449  
QY 1505 GAAGCAAGAGAGAGAGGAGGAAAGTAACAGAGAGGTGCGTAGTACACACGAGGTTGAAG 1564  
DB 450 GluAspArgLysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSer 467  
QY 1565 GAAGGCGATGCTTTCATCATGCCAGCTCATCCAGTAGCCATCAAGCTTCTCTCCGAA 1624  
DB 468 ProGlyAspValIleIleProAlaGlyHisProValAlaIleSerAlaSerAsn 487  
QY 1625 CTCCATCTCTGCTGCTTCCGATCAACGCTGAAAAACAACACAGAACTTCTCTTCGAGGT 1684  
DB 488 LeuAsnLeuLeuGlyPheGlyIleAsnAlaLysAsnAsnGlnArgAsnPheLeuSerGly 507  
QY 1685 GATAGGACAATGTGATAGACCATAGAGAGAGCGAAGGATTTAGCATTCCTCTGGG 1744  
DB 508 SerAspAspAsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGly 527  
QY 1745 TCGGGTGAACAAGTTGAGAGCTCATCAAAAACCAAGAGGAATCTCACCTTGTGAGTGCT 1804  
DB 528 SerSerGlnGluValAsnArgLeuIleLysAsnGlnLysGlnSerHisPheAlaSerAla 547  
QY 1805 CGTCTCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAAAAGAGTCTCTCTGAGAAAGAG 1864  
DB 548 GluPro-----Glu 550  
QY 1865 GATCAAGAGGAGGAAAACCAAGGAGGAGGTCCACTCTCTTCAATTTTGAAGGCTTTT 1924



```
Db 551 GlnIysGluGluGluSerGlnArgLysArgSerProLeuSerSerValLeuAspSerPhe 570
RESULT 13
Q948X9 PRELIMINARY; PRT; 623 AA.
AC Q948X9:
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-conglycinin alpha-subunit.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Fukazawa C.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030839; BAB64304.1; -.
DR HSSP; P25974; 1IPJ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmC_Like_cupin.
DR Pfam; PF00190; Cupin; 2.
SQ SEQUENCE 623 AA; 72475 MW; EBF033DA1316AE9A CRC64;

Alignment Scores:
Pred. No.: 1-28e-74 Length: 623
Score: 1265.50 Matches: 283
Percent Similarity: 58.22% Conservative: 110
Best Local Similarity: 41.93% Mismatches: 179
Query Match: 35.53% Indels: 103
DB: 2 Gaps: 18

US-10-728-051-1 (1-2032) x Q948X9 (1-623)
Qy 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGATCCTGTGCTGGCTTCAGTT 109
Db 1 MetArgAlaArgPhePro-----LeuLeuLeuGlyValPheLeuAlaSerVal 18
Qy 110 TGTGCAACGATGCCAAGTCATCCCTACAGAGAGAAACAGAGAACCCCTGCGCCAG 169
Db 19 SerValSerPheGlyIleAla-----TyrTrpGluLys---GlnAsnProLysHisAsn 35
Qy 170 AGTGCTCCAGAGTTCTCAACAGAACCGGATGCTTGAAGCAAAAGGATCGAGTCT 229
Db 36 LysCysAsnLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 55
Qy 230 CGTGTGACAC-----AAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACAC 283
Db 56 ArgCysAsnLeuLeuLysValGluLysGluGluCys-----GluGluGlyGlu 72
Qy 284 ACTGGCAACCAACCAACGTTTCCCTCCAGGGAGGCGGACACGTGCGCCCAACCCGA 343
Db 73 IleProArgProArgProArgProGlnHisProGluArg---GluProGlnGlnProGly 91
Qy 344 GACTACCATGATGACCC-----CGTCAA 367
Db 92 GluLysGluGluAspGluAspGluGlnProArgProPheProArgProArgGln 111
Qy 368 CCCGAGAGAGAGAA----- 382
Db 112 ProArgGlnGluGluGluHisGluGlnArgGluGluGluTrpProArgLysGluGlu 131
Qy 383 ---GGAGCCGATGGGA-----CCAGCT 403
Db 132 LysArgGlyGluLysGlySerGluGluGluGlnAspGlyArgGluHisProArgProHis 151
Qy 404 GGACCGAGGAGCGTGAAGAGAGAGAGACTGGAGACAACCAAGAGAGATTGGAGCGGA 463
Db 152 GlnProHisAspGluAspGluGluGlnAspGluGln-----PheProPheProArg 169
```

```
Qy 464 CCAAGTCAATCAGCAGCCACGGAATAAGG-----CCGAA 499
Db 170 ProHisGlnLysGluSerGluGluArgLysGlnGluAspGluAspGluGln 189
Qy 500 GGAGAGAGAGAGAACAGAGTGGGNAACACAGGTAGCATGTAGGGAAGAAACATCT 559
Db 190 GlnArgGluSerGluGluSerGluSerGlnArgGluLeuArgArgHisLys 209
Qy 560 CGAAACAACCCCTTCTACTTCCCGTCAAGCGCGTTTACACCCGCTACCGGAACCAAAAC 619
Db 210 AsnLysAsnProPheHisPheGlySerAsnArgPheGluThrLeuPheLysAsnGlnTyr 229
Qy 620 GTTAGATCCGGTCTCTGAGAGGTTTGACCAAGGTCAGAGCAGTTTCAAGATCTCCAG 679
Db 230 GlyArgIleArgValLeuGlnArgPheAsnGlnArgSerProGlnLeuGlnAsnLeuArg 249
Qy 680 AATCAGCTATTGTGCAGATCGAGGCGCAACCTAACACTCTTCTTCTCCNAGCAGCT 739
Db 250 AspTyrArgIleLeuGluPheAsnSerLysProAsnThrLeuLeuLeuProAsnHisAla 269
Qy 740 GATGCTGATAACATCTTGTATTCCAGCAAGGGCAAGCCACCGTAGCAATATGGC 799
Db 270 AspAlaAspTyrLeuIleAlaIleLeuAsnGlyThrAlaIleLeuSerLeuValAsnAsn 289
Qy 800 AATAACAGAAAGAGCTTTAATTTGACGAGGCGCATCTCAGATCCAGATCCCGTTTC 859
Db 290 AspAspArgAspSerTyrArgLeuGlnSerGlyAspAlaLeuArgValProSerGlyThr 309
Qy 860 ATTCTCTACATCTTGAAACCGCCATGACACCAAGAACTCAGATAGTATAAATCTCCATG 919
Db 310 ThrTyrTyrValValAsnProAspAsnAsnGluAsnLeuLeuThrLeuAlaIle 329
Qy 920 CCGTTTAAACACACCGCGGATTTGAGGATTTCTCCCGCGAGCAGCCGAGACCAATCA 979
Db 330 ProValAsnLysProGlyArgPheGluSerPhePheLeuSerSerThrGluAlaGlnGln 349
Qy 980 TCTACTTCAGGGCTTCAGAGAAATACGTTCGAGCGCCCTTCATTCGGAATTCAT 1039
Db 350 SerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAlaSerTyrAspThrLysPheGlu 369
Qy 1040 GAGATCAGGAGGTGCTGTAGAGAGATGACAGGAGGTGACGAGGAGGAGAGGCGCAG 1099
Db 370 GluIleAsnLysValLeuPheSerArgGluGluGly-----GlnGlnGlnGlyGlu 386
Qy 1100 AGCGATGAGTACTCGAGTACTGAGAACAAATGAAGGAGTGTATGATCAAGTGTCAAAG 1159
Db 387 GlnArgLeu-----GlnGluSerValIleValGluIleSerLys 399
Qy 1160 GAGCAGCTTGAAGAACTTACTAAGCAGCTTAATCCGTCTCAAGAAAGAGCTCCGAAGAA 1219
Db 400 GluGlnIleArgAlaLeuSerLysArgAlaLysSerSerSerArgLysThrIleSerSer 419
Qy 1220 GAGGAGATATCACCAACCCCACTTACAGAGAGGCGCGCCGATCTTTCTAACAC 1279
Db 420 Glu-----AspLysProPheAsnLeuArgSerArgAspProIleTyrSerAsnLys 436
Qy 1280 TTTGGGAAGTTATTTCAGGTGAAGCAGACAAAGAACCCCGAGCTTCAGGACCTGCAC 1339
Db 437 LeuGlyLysPhePheGluIleThrProGlu---LysAsnProGlnLeuArgAspLeuAsp 455
Qy 1340 ATGATCTCCTACCTGTGTAGAGATCAAGAGAGAGCTTTGATGTCTCCACACTTCACTCA 1399
Db 456 IlePheLeuSerIleValAspMetAsnGluGlyAlaLeuLeuLeuProHisPheAsnSer 475
Qy 1400 AAGCCATCGTTATCTGCTCGTCAACAAGAGAACTGGAAACCTTGNACTCGGTGTA 1459
Db 476 LysAlaIleValIleLeuValIleAsnGluGlyAspAlaAsnIleGluLeuValGlyLeu 495
Qy 1460 AGAAAGAGCAACACAGAGGGGACCGCGGGAAGAGAGGAGCAGAAAGAGAGAGAG 1519
Db 496 ---LysGluGlnGln-----GluGluGlnGln 504
```

```

QY 1520 GAGGAAGTAACAGAGAGGTGCTAGGTATACACAGCGAGGTGTAAGGAAGCGCATGTGTC 1579
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
505 GluGluGlnProLeuGluValAlaGlyLysTyrArgAlaGluLeuSerGluGlnAspIlePhe 524
QY 1580 ATCATGCCAGCAGCTCATCAGTAGCCATCAACGCTTCCTCCGAATCCTCATCTGCTTGGC 1639
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
525 ValIleProAlaGlyTyrProValValAlaAsnAlaThrSerAsnLeuAsnPheAla 544
QY 1640 TTCGGTATCAACCTGAAACACACAGATCTCTCTTGCAGGTGATTAAGGACAAATGTG 1699
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
545 IleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeuAlaGlySerGlnAspAsnVal 564
QY 1700 ATAGACCATAGAGAAGCAAGCAAGATTTAGCATTCCTCGGTGCGGTGCAACAAGTT 1759
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
565 IleSerGlnIleProSerGlnValGlnGluLeuAlaPheLeuGlySerAlaGlnAlaVal 584
QY 1760 GAGAAGCTCATCAAAACCAAGAGGAATCTCATTGTTGTGAGTCTGCTCTCAATCTCAA 1819
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
585 GluLysLeuLeuLysAsnGlnArgGlySerTyrPheValAspAlaGlnProLysLys--- 603
QY 1820 TCTCAATCTCCGTCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAAGAGGAGAA 1879
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
604 -----LysGluGluGly 607

QY 1880 AACCAAGAGGGAAGGTCCACTCTTCAATTTGAAGCTTTT 1924
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
608 AsnLysGlyArgLysGlyProLeuSerSerIleLeuArgAlaPhe 622

RESULT 14
Q41674 Q41674 PRELIMINARY; PRT; 545 AA.
AC Q41674;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Convicillin precursor.
OS Vicia narbonensis (Narbonne vetch).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Vicia.
OX NCBI_TaxID=3912;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cotyledon;
RA Koch G., Koenig S., Becker C., Horstmann C., Schlesier B.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z71986; CAA96513.1; -.
DR HSP; P25974; 11PJ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmlC_Ilike_cupin.
DR Pfam; PF00190; Cupin; 2.
KW Signal.
FT SIGNAL 1 28 Potential.
FT CHAIN 29 545 7S globulin; convicillin.
SQ SEQUENCE 545 AA; 62810 MW; 459A876F92F5A87E CRC64;

Alignment Scores:
Pred. No.: 3,56e-74 Length: 545
Score: 1259.50 Matches: 279
Percent Similarity: 58.18% Conservative: 98
Best Local Similarity: 43.06% Mismatches: 150
Query Match: 35.33% Indels: 121
DB: 2 Gaps: 16

US-10-728-051-1 (1-2032) x Q41674 (1-545)
QY 38 GTACGAGCAATGACAGGAGGAGGTTCTCCACTGATGCTGTGCTAGGATCTTGTC 97
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MetAlaThrAlaMetLysSerArgPhe---ProValLeuLeuLeuGlyIleIle 19
QY 98 CTGCTTCAGTTTCTGCAACGCATGCCAAGTCATCCTTACCAGAAAGAAACAGAGAAC 157

```

```

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
20 LeuAlaSerLeuCysValThrTyrAlaAsn-----TyrAspGluGlyThrGlu--- 35
QY 158 CCCTCGCGCCAGAGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAG 217
Db ----- 35
QY 218 GCATCGAGTCTCGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCTCTGA 277
Db -----ProArgVal----- 38
QY 278 GGACACACTGGCACACCAACCAACGTTCCCTCCAGGGAGCGGACACGTCGGCGCCAA 337
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39 -----ProGlyGlnArgGluArgGlyArgGln 47
QY 338 CCGGAGACTACGATGATGACCGCGTCAACCCCGAAGAGAGAGGCGCGATGGGGA 397
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 GluGlyGluLysGluGluLysArgHis-----GlyGluTrpArg 60
QY 398 CCA-----GCTGGACCGAGGAGCGTGAAAGA----- 424
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 ProSerHisGluLysGluAlaGlnProGlyArgArgGluArgTrpGluThrSerGluGlu 80
QY 425 -----GAAGAAGACTGGAGA-----CAACCAAGAGAGATTGGAGGCGACA 466
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 GluGluArgValAspGluGluTrpArgGlySerGlnArgHisGluAsp-----Pro 97
QY 467 AGTCATCAGACCCACCGGAAATAAGCCCGCAAGGAAGAGAGAGGAGAAACAAGAGTGG--- 523
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 GluGluArgAlaArgGluArgTyrArgAlaGluGluArgGluArgArgGlnTrpGlu 117
QY 524 GGAACACACGAGTACCATGTGAGGAAGAAACATCTCGGAACAACCTTCTACTTCCG 583
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
584 TCAAGCGCGTTTAGCACCGCTACCGGAACCAACGCTAGCATCCGCTCTCGAGAGG 643
QY 138 SerAsnLysPheLeuThrLeuPheGluAsnGluAsnGlyHisIleArgArgLeuGlnArg 157
QY 644 TTTGACCAAGGTCGAAGGCGAGTTTCAGAAATCTCCAGAATCACCGTATTGTGCGAGTCGAG 703
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
158 PheAspLysArgSerAspLeuPheGluAsnLeuGlnAsnTyrArgLeuValGluTrpArg 177
QY 704 GCAAACTCAACATCTCTGTTCTTCCCAAGCAGCGTGATGCTGATAAATCTTCTTATC 763
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
178 AlaLysProHisThrIlePheLeuProGlnHisIleAspAlaAspLeuIleLeuThrVal 197
QY 764 CAGCAAGGCAAGCCACCGTACCGCTAGCAATGCAATTAACAGAAAGAGCTTTAATCTT 823
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 LeuSerGlyArgAlaIleLeuThrValLeuSerProAsnAspArgAsnSerTyrAsnLeu 217
QY 824 GACGAGGCGCATGCAATCAGAAATCCCATCCGTTTTCATTTCTTACATCTTGAACCGCAT 883
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
218 GluArgGlyAspThrIleLysLeuProAlaGlyThrThrSerTyrLeuLeuAsnGlnAsp 237
QY 884 GACAACCAACCTCAGAGTAGCTAAATCTCCATGCCGTTAAACACACCCCGCCAGTTT 943
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
238 AspGluGluAspLeuArgValValAspLeuSerIleSerValAsnArgProGlyLysVal 257
QY 944 GAGGATTTCTTCCCGCGCAGCAGCGAGACCAATCATCTTCTGAGCGGCTTCAGCAGG 1003
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
258 GluSerPheGlyLeuSerGlySerLysAsnGln-----TyrLeuArgGlyPheSerLys 275
QY 1004 AATACCTTGGAGCGCCCTTCAATTCGGAAATTCATAGATACGAGGCGGTGCTGTAGAA 1063
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
276 AsnIleLeuGluAlaSerLeuAsnThrLysTyrGluThrIleGluLysValLeuLeuGlu 295
QY 1064 GGAATGCGAGGAGTACGACAGGAGAGGAGCGGCGATGAGTACTCGAGTACT 1123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
296 Glu-----ProGlnGlnSerIleGlyGlnLysArgSerGlnArgGlnGlu 311
QY 1124 GAGAACAAATGAAGAGTGTAGTCAAGTGTCAAAGGAGCACGTTGAAGAACTTACTAG 1183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

312 ThrAsn-----AlaLeuValLysValSerHrgGluInValGluLeuLysArg 328  
1184 CACGCTAAATCCGTCTCAAGAAAGGCTCCGAGAAGAGGAGATATACCAACCCCAATC 1243  
|||||  
329 LeuAlaLysSerSerLysLysGlyValSerSerGluPheGlu-----ProPhe 345  
|||||  
1244 AACTTGAGAGNAGCGGCGCCGCTCTTCTAAACACTTTGGGAAGTTATTGTAGGTGAAG 1303  
|||||  
346 AsnLeuArgSerGlnAsnProLysTyrSerAsnLysPheGlyLysLeuPheGluLeuThr 365  
|||||  
1304 CCAGACAAGAAAGAACCCCGCTTCAGACCTCGACATGATGCTCACCTGTGTAGAGATC 1363  
|||||  
366 ProGluLysLysTyrProGlnLeuGlnAspLeuAspIlePheValSerValGluLeu 385  
|||||  
1364 AAGAAGAGGAGCTTTGATGCTCCACACTCAACTCAAAGGCCATGTGTTATCGTCGTCGT 1423  
|||||  
386 AsnGluGlyGlyLeuMetLeuProHisTyrAsnSerArgAlaIleValIleLeuLeuVal 405  
|||||  
1424 AACAAAGGAAGCTGAACCTTGAACCTCGTGGCTGTAAAGAAAGCAACAAACAGAGGGGA 1483  
|||||  
406 AsnGluGlyLysGlyAsnLeuGluLeuValGlyLeuLysAsnGluGlnGluGlnArg 425  
|||||  
1484 CGGCGGCAAGAAAGAGGAGGACGAAGACGAAGAGGAGGGAAGTACAGAGAGGTGCGT 1543  
|||||  
426 GluArgGlnAspGluGln-----GlnValGln 434  
|||||  
1544 AGGTACACAGCAGGAGTTGAAGAGGAGCGGCTGTGTTTCATCATGCCAGAGCTCATCCAGTA 1603  
|||||  
435 ArgTyrGluAlaArgLeuSerProGlyAspValIleIleProIaGlyHisProVal 454  
|||||  
1604 GCCATCAACGCTTCTCCGAACTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAAC 1663  
|||||  
455 AlaValSerAlaSerAsnLeuAsnLeuLeuGlyPheGlyIleAsnAlaGluAsnAsn 474  
|||||  
1664 CACAGATCTTCCTTCGAGCTGATAAGCAATGTGTATACCCAGATAGAGAAGCAAGCG 1723  
|||||  
475 GlnArgAsnPheLeuThrGlySerAspAspValIleSerGlnIleGluAsnProVal 494  
|||||  
1724 AAGGATTATGACATTCCCTCGGTCCGGTGAACAAAGTTGAGAAGGCTCATCAAAAACCAAG 1783  
|||||  
495 LysGluLeuThrPheProGlySerAlaGlnGluValAsnArgLeuLeuLysAsnGlnGlu 514  
|||||  
1784 GAATCTCACTTTGTAGTGCTCGTCTCAATCTCAATCTCAATCTCGTCTGCTCTGAG 1843  
|||||  
515 HisSerHisPheAlaAsnAlaGluPro----- 523  
|||||  
1844 AAGAGTCTCTGAGAAGAGGATCAAGACGAGGAGGAAACCAAGAGGAGGAGGTCCACTC 1903  
|||||  
524 -----GluGlnLysGlyGluSerGlnArgLysArgSerProIle 537  
|||||  
1904 CTTTCAATTTGAAGGCTTTTAAAC 1927  
|||||  
538 SerSerIleLeuGlyThrPheAsn 545  
|||||  
  
RESULT 15  
Q9W3X6 PRELIMINARY; PRT; 613 AA.  
ID Q9W3X6  
AC Q9W3X6;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Convicilin precursor.  
GN Name=cvc;  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum. ...  
NCBI\_TaxID=3888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Seed cotyledon;  
RA Casey R.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 8, 2005, 04:11:27 ; Search time 19.3806 Seconds  
(without alignments)  
5523.384 Million cell updates/sec

Title: US-10-728-051-2

Perfect score: 1265

Sequence: 1 gctcaccatactagtagcccc.....taaaaagatcatgttttggtt 717

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying Chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO.spool_p/US10728051/runat_07042005_125046_14072/app.query.fasta_1.4757
-DB=Issued Patents AA -QFMT=faстан -SUFFIX=rai -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=BITS -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728051 @CGN 1 1 122 @runat_07042005_125046_14072 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	842	66.6	157	4	US-09-106-872A-22
2	842	66.6	157	4	US-09-191-593-10
3	842	66.6	159	4	US-09-106-872A-2
4	827	65.4	157	4	US-09-191-593-21
5	243	19.2	158	2	US-08-618-911-4
6	243	19.2	158	3	US-08-938-675A-2
7	243	19.2	158	4	US-09-531-727-2
8	241.5	19.1	155	2	US-08-618-911-2
9	229	18.1	158	2	US-08-618-911-6
10	156	12.3	28	4	US-09-191-593-53
11	156	12.3	158	2	US-08-670-186-4
12	151	11.9	158	2	US-08-670-186-6

13	139	11.0	155	2	US-08-670-186-2	Sequence 2, Appli
14	139	11.0	184	1	US-08-453-924-3	Sequence 3, Appli
15	138	10.9	28	4	US-09-191-593-58	Sequence 58, Appl
16	136.5	10.8	164	4	US-09-627-536-13	Sequence 13, Appl
17	118	9.3	30	4	US-09-106-872A-16	Sequence 16, Appl
18	118	9.3	30	4	US-09-191-593-57	Sequence 57, Appl
19	106.5	8.4	174	4	US-09-645-593-7	Sequence 7, Appli
20	106.5	8.4	1088	4	US-09-233-857-13	Sequence 13, Appl
21	105.5	8.3	1099	3	US-09-442-100-2	Sequence 2, Appli
22	105.5	8.3	1099	4	US-08-939-106-2	Sequence 2, Appli
23	105.5	8.3	1099	4	US-09-442-102-2	Sequence 2, Appli
24	102	8.1	316	3	US-08-098-3278-31	Sequence 31, Appli
25	102	8.1	316	3	US-08-462-625-31	Sequence 31, Appl
26	100	7.9	17	4	US-09-191-593-55	Sequence 55, Appl
27	100	7.9	17	4	US-09-191-593-61	Sequence 61, Appl
28	100	7.9	17	4	US-09-191-593-62	Sequence 62, Appl
29	100	7.8	224	4	US-09-270-767-44901	Sequence 44901, A
30	99.5	7.9	148	4	US-09-083-852A-4	Sequence 4, Appli
31	99.5	7.9	148	4	US-09-083-852A-6	Sequence 6, Appli
32	99.5	7.9	148	4	US-09-489-674B-6	Sequence 6, Appli
33	97	7.7	185	4	US-09-252-991A-23490	Sequence 23490, A
34	95	7.5	32	1	US-08-158-704-1	Sequence 1, Appli
35	95	7.5	32	2	US-08-610-424B-1	Sequence 1, Appli
36	95	7.5	32	4	US-09-336-463-1	Sequence 1, Appli
37	95	7.5	32	4	US-09-191-593-1	Sequence 1, Appli
38	95	7.5	152	4	US-09-252-991A-19284	Sequence 19284, A
39	94	7.4	43	3	US-09-303-814-1	Sequence 1, Appli
40	94	7.4	500	4	US-09-265-630-13	Sequence 13, Appl
41	93.5	7.4	415	4	US-09-252-991A-18252	Sequence 18252, A
42	93	7.4	537	4	US-09-252-991A-27024	Sequence 27024, A
43	92.5	7.3	788	2	US-08-918-914-4	Sequence 4, Appli
44	91.5	7.2	1898	1	US-08-056-200-94	Sequence 94, Appl
45	91.5	7.2	1898	2	US-08-800-644-94	Sequence 94, Appl

#### ALIGNMENTS

RESULT 1  
US-09-106-872A-22  
; Sequence 22, Application US/09106872A  
; Patent No. 6486311  
; GENERAL INFORMATION:  
; APPLICANT: Burks Jr., A. Wesley  
; APPLICANT: Stanley, J. Steven  
; APPLICANT: Cockrell, Gael  
; APPLICANT: King, Nina E.  
; APPLICANT: Sampson, Hugh A.  
; APPLICANT: Helm, Ricki M.  
; APPLICANT: Bannan, Gary A.  
; TITLE OF INVENTION: Peanut Allergens and Methods  
; FILE REFERENCE: HS 103 CIP  
; CURRENT APPLICATION NUMBER: US/09/106,872A  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: PCT/US96/15222  
; PRIOR FILING DATE: 1996-09-23  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 157  
; TYPE: PRT  
; ORGANISM: Arachis hypogaea  
US-09-106-872A-22

Alignment Scores:  
Pred. No.: 8.73e-86 Length: 157  
Score: 842.00 Matches: 157  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 66.56% Indels: 0  
DB: 4 Gaps: 0

US-10-728-051-2 (1-717) x US-09-106-872A-22 (1-157)

Qy	2	CTCACCATCTAGTAGCCCTCGCCCTTTCTCTCGTGCCTCCACGATCTCGGAGCGAG	61
Db	1	LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAhiAlaSerAlaLeuArgIn	20
Qy	62	CAGTGGGAACCTCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTTGAGG	121
Db	21	GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg	40
Qy	122	CCCTGCGAGCAACATCTCATTCAGAGAAGATCCAACGTCACGAGGATTCATATGAACCGGAC	181
Db	41	ProCysGluGlnHisLeuMetGlnIlyIleGlnArgAspGluAspSerTyrGluArgAsp	60
Qy	182	CCGTACAGCCCTAGTCAGAGATCCGTACAGCCCTAGTCCATATGATCCGAGAGGCGCTGGA	241
Db	61	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly	80
Qy	242	TCCTCTCAGCACCAAGAGAGGTGTTCGAATCAGCTGAACAGATTTGAGAACCAACCAAGG	301
Db	81	SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg	100
Qy	302	TGATGTGGAGGCATTGCAACAGATCATGAGAACCCAGAGCGATAGTTGCAGGGGAGG	361
Db	101	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	120
Qy	362	CAACAGGAGCAACAGTTCAAGAGGAGAGCTCAGGAACCTTGCCCTCAACAGTCGGCGCTTAGG	421
Db	121	GlnGlnGlnGlnPheIlyshArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg	140
Qy	422	GCACCAACAGCTTGCGACTTGCACTCGCAAGTGGCGGCGAGAGACAGATAC	472
Db	141	AlaProGlnArgCysAspLeuAspValGluSerGlyArgGlyArgAspArgTyr	157

## RESULT 2

US-09-191-593-10

US-09-191-993-10  
; Sequence 10, Application US/09191593

; Patent No. 6835824

FACEID NO. 0835824  
: GENERAL INFORMATION:

APPLICANT: BURKS, A Wesley, HELM, Ricki M,

APPLICANT: COCKRELL, Gael, STANLEY, J Steven,

APPLICANT: BANNON, Gary A

TITLE OF INVENTION: PEANUT ALLERGENS AND

INVENTOR:	TITLE OF INVENTION:	METHODS:
...	...	...

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESSEE: Head, Johnson & Kachigian

STREET: 112 W. Center St., Suite 230

; CITY: Fayetteville

STATE: Arkansas AR

; COUNTRY: United States of America

; ZIP: 72701

; COMPUTER READABLE FORM:

```

;
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb

```

;  
MEDIUM TYPE: storage

**COMPUTER:** IBM PC compatible

```
; OPERATING SYSTEM: MS-DOS 6.2
```

```

; SOFTWARE: Wordperfect 6.0C
;

```

```

; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/09/191,593  
FILING DATE: 12 NOVEMBER 1999

FILING DATE: 13 NOVEMBER 1998

CLASSIFICATION:

; PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER.

APPLICATION NUMBER: 08/717,933

FILING DATE: 23 SEPTEMBER 1996  
APPLICATION NUMBER: IIS 07/998

APPLICATION NUMBER: US 07/998,377  
FILING DATE: 20 DECEMBER 1992

FILING DATE: 30 DECEMBER 1992  
APPLICATION NUMBER: IIS 08/158 704

APPLICATION NUMBER: US 08/158,704  
FILING DATE: 28 NOVEMBER 1993

FILING DATE: 29 NOVEMBER 1993  
APPLICATION NUMBER: IIS 60/009 455

APPLICATION NUMBER: US 60/009,455  
FILING DATE: 29 DECEMBER 1995

;  
 FILING DATE: 29 DECEMBER 1995  
 ;  
 APPLICATION NUMBER: IIS 08/610 424

APPLICATION NUMBER: US 08/610,424  
FILING DATE: 04 MARCH 1996

ATTORNEY/AGENT INFORMATION:  
FILING DATE: 04 MARCH 1996

ATTORNEY/AGENT INFORMATION:  
NAME: ALEXANDER DANIEL

NAME: ALEXANDER, DANIEL R

;	REGISTRATION NUMBER: 32,604	
;	REFERENCE/DOCKET NUMBER: ARK00895601B	
;	TELECOMMUNICATION INFORMATION:	
;	TELEPHONE: (501) 582-9111	
;	TELEFAX: (501) 521-4931	
;	TELEX: No. 6835824 applicable	
;	INFORMATION FOR SEQ ID NO: 10:	
;	SEQUENCE CHARACTERISTICS:	
;	LENGTH: 157 amino acids	
;	TYPE: amino acid	
;	STRANDEDNESS: not applicable	
;	TOPOLOGY: Unknown	
;	MOLECULE TYPE: cDNA	
;	DESCRIPTION: identified as Ara h II cDNA clone	
;	DESCRIPTION: derived amino acid sequence	
;	HYPOTHETICAL: No	
;	ANTI-SENSE: No	
;	FRAGMENT TYPE: No. 6835824 applicable	
;	ORIGINAL SOURCE:	
;	ORGANISM: Arachis hypogaea	
;	STRAIN: Florunner	
;	INDIVIDUAL ISOLATE:	
;	DEVELOPMENTAL STAGE: seed	
;	HAPLOTYPE: No. 6835824 applicable	
;	TISSUE TYPE: seed cDNA	
;	CELL TYPE: No. 6835824 applicable	
;	CELL LINE: No. 6835824 applicable	
;	ORGANELLE: No. 6835824 applicable	
;	IMMEDIATE SOURCE:	
;	LIBRARY:	
;	CLONE:	
;	POSITION IN GENOME:	
;	CHROMOSOME/SEGMENT: No. 6835824 applicable	
;	MAP POSITION: No. 6835824 applicable	
;	UNITS: No. 6835824 applicable	
;	FEATURE:	
;	NAME/KEY:	
;	LOCATION:	
;	IDENTIFICATION METHOD: By agreement with	
;	IDENTIFICATION METHOD: protein information and established	
;	IDENTIFICATION METHOD: consensus sequence	
;	OTHER INFORMATION: Seed storage protein and	
;	OTHER INFORMATION: allergen	
;	PUBLICATION INFORMATION:	
;	AUTHORS:	
;	TITLE:	
;	JOURNAL:	
;	VOLUME:	
;	ISSUE:	
;	PAGES:	
;	DATE:	
;	DOCUMENT NUMBER:	
;	FILING DATE:	
;	PUBLICATION DATE:	
;	RELEVANT RESIDUES IN SEQ ID NO:	
;	US-09-191-593-10	
;	Alignment Scores:	
;	Pred. No.: 8,73e-86	Length: 157
;	Score: 842.00	Matches: 157
;	Percent Similarity: 100.00%	Conservative: 0
;	Best Local Similarity: 100.00%	Mismatches: 0
;	Query Match: 66.56%	Indels: 0
;	DB: 4	Gaps: 0
;	US-10-728-051-2 (1-717) x US-09-191-593-10 (1-157)	
Qy	2 CTCACCATAGTAGCCCTCGCCCTTTTCCTCCTCGTCCCAAGCAGTCTG	
Db	1 LeuthrriLeuValaLeuValaLeuPheLeuValaAlaHisAlaSerA	
Oy	62 CAGTGGGAACCTCCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGGCG	



```
Db 21 GlnTrrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGGAGCAACATCTCAGCAGAGATCCACGTCAGCAGGATTCATATGACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATCATCGAGAGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGATGTCCAGGAGGATTCGAACAGATCATCGAGAGCAACAGAGCGATAGTTGCGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTSCGGCTTAGG 421
Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAGAGGTTGCGACTTGGACGTGAAAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 3
US-09-106-872A-2
; Sequence 2, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannon, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Arachis hypogaea
; FEATURE:
; OTHER INFORMATION: Amino Acids 17-26 are Ara H 2 Binding Epitopes,
; OTHER INFORMATION: Peptide 1
; OTHER INFORMATION: Amino Acids 23-32 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 2
; OTHER INFORMATION: Amino Acids 29-38 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 3
; OTHER INFORMATION: Amino Acids 41-50 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 4
; OTHER INFORMATION: Amino Acids 51-60 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 5
; OTHER INFORMATION: Amino Acids 59-68 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 6
; OTHER INFORMATION: Amino Acids 67-76 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 7
; OTHER INFORMATION: Amino Acids 117-126 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 8
; OTHER INFORMATION: Amino Acids 129-138 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 9
; OTHER INFORMATION: Amino Acids 145-154 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 10
; OTHER INFORMATION: Amino Acids 13-22 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 1, Table 12
```

```
; OTHER INFORMATION: Amino Acids 27-36 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 2, Table 12
; OTHER INFORMATION: Amino Acids 59-68 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 3, Table 12
; OTHER INFORMATION: Amino Acids 81-90 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 4, Table 12
; OTHER INFORMATION: Amino Acids 91-100 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 5, Table 12
; OTHER INFORMATION: Amino Acids 145-154 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 6, Table 12
US-09-106-872A-2
Alignment Scores:
Pred. No.: 8,77e-86 Length: 159
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 4 Gaps: 0
US-10-728-051-2 (1-717) x US-09-106-872A-2 (1-159)
QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTTCCTCTCCTCGCTGCCACGCATCTCGAGGCGAG 61
Db 3 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 22
QY 62 CAGTGGAACTCCAGGAGACAGAGATGCCAGAGCAGCTCGAGAGGGGGAACCTGAGG 121
Db 23 GlnTrrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 42
QY 122 CCCTGGAGCAACATCTCATGCAAGATCCACGTCACACGTCGAGGATTCATATGAACGGAC 181
Db 43 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 62
QY 182 CCGTCAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATCATCGAGAGCGCTGGA 241
Db 63 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 82
QY 242 TCCTCTCAGCACCAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301
Db 83 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 102
QY 302 TGATGTCCAGGAGGATTCGAACAGATCATCGAGAGCAACAGAGCGATAGTTGCGGGAGG 361
Db 103 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 122
QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTSCGGCTTAGG 421
Db 123 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 142
QY 422 GCACCAGAGGTTGCGACTTGGACGTGAAAGTGGCGGCGAGACAGATAC 472
Db 143 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 159

RESULT 4
US-09-191-593-21
; Sequence 21, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS 6.2  
SOFTWARE: Wordperfect 6.0C  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/191,593  
FILING DATE: 13 NOVEMBER 1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/717,933  
FILING DATE: 23 SEPTEMBER 1996  
APPLICATION NUMBER: US 07/998,377  
FILING DATE: 30 DECEMBER 1992  
APPLICATION NUMBER: US 08/158,704  
FILING DATE: 29 NOVEMBER 1993  
APPLICATION NUMBER: US 60/009,455  
FILING DATE: 29 DECEMBER 1995  
APPLICATION NUMBER: US 08/610,424  
FILING DATE: 04 MARCH 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: ALEXANDER, DANIEL R  
REGISTRATION NUMBER: 32,604  
REFERENCE/DOCKET NUMBER: ARK00895601B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (501) 582-9111  
TELEFAX: (501) 521-4931  
TELEX: No. 6835824 applicable  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 157 amino acids  
TYPE: amino acid  
STRANDEDNESS: No. 6835824 applicable  
TOPOLOGY: unknown  
MOLECULE TYPE: glycoprotein  
DESCRIPTION: identified as Ara h 2 P38 deduced  
DESCRIPTION: sequence from nucleotide sequence reading frame.  
HYPOTHETICAL: No  
ANTI-SENSE: No. 6835824 applicable  
FRAGMENT TYPE: No. 6835824 applicable  
ORIGINAL SOURCE:  
ORGANISM: Arachis hypogaea  
STRAIN: Florunner  
INDIVIDUAL ISOLATE: Ara h II  
DEVELOPMENTAL STAGE:  
HAPLOTYPE: No. 6835824 applicable  
TISSUE TYPE:  
CELL TYPE: No. 6835824 applicable  
CELL LINE: No. 6835824 applicable  
ORGANELLE: No. 6835824 applicable  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: No. 6835824 applicable  
MAP POSITION: No. 6835824 applicable  
UNITS: No. 6835824 applicable  
FEATURE:  
NAME/KEY: final check  
LOCATION: 4..480  
IDENTIFICATION METHOD: By agreement with  
IDENTIFICATION METHOD: protein information and established  
IDENTIFICATION METHOD: consensus sequence  
OTHER INFORMATION: Seed storage protein and  
OTHER INFORMATION: allergen  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:

DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-09-191-593-21  
Alignment Scores:  
Pred. No.: 4.19e-84 Length: 157  
Score: 827.00 Matches: 155  
Percent Similarity: 98.73% Conservative: 0  
Best Local Similarity: 98.73% Mismatches: 2  
Query Match: 65.38% Indels: 0  
DB: 4 Gaps: 0  
US-10-728-051-2 (1-717) x US-09-191-593-21 (1-157)  
QY 2 CTCACCATAGTAGCCCTCGCCCTTTTCCTCGCTGCCGACGATCTCGGAGGAG 61  
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20  
QY 62 CAGTGGAACTCCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121  
Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40  
QY 122 CCCTCGAGCAACATCTCATGCAAGATCCAACTGACGAGGATTCATATGAACGGGAC 181  
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60  
QY 182 CGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGA 241  
Db 61 ProSerTyrProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80  
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGACAGCTTTGAGAACCAACCAAGG 301  
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100  
QY 302 TGCATGTGGAGGCAATTCACAGATCATATGAGAACCAACGAGGAGTGTGAGGGGAGG 361  
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120  
QY 362 CAACAGGAGCAACAGTTCAGAGAGGGAGGCTCAGAACTTGCCTCAACAGTGGCGCTTAGG 421  
Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140  
QY 422 GCACCACAGCGTTGGGACTTGGACGTTCGAAAGTGGCGGAGAGACAGATAC 472  
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157  
RESULT 5  
US-08-618-911-4  
Sequence 4, Application US/08618911  
Patent No. 5850016  
GENERAL INFORMATION:  
APPLICANT: Jung, Rudolf  
APPLICANT: Hastings, Craig  
APPLICANT: Coughlan, Sean  
APPLICANT: Hu, David  
TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN  
TITLE OF INVENTION: SEEDS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.  
STREET: 700 Capital Square, 400 Locust Street  
CITY: Des Moines  
STATE: Iowa  
COUNTRY: USA  
ZIP: 50309  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:



GENERAL INFORMATION:  
APPLICANT: de Lumen, Benito O.  
Galvez, Alfredo F.  
TITLE OF INVENTION: Lunasin Peptides  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 75 DENISE DRIVE  
CITY: HILLSBOROUGH  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94010  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/531,727  
FILING DATE: 21-Mar-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/938,675  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B98-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 343-4341  
TELEFAX: (650) 343-4342  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-531-727-2

Alignment Scores:  
Pred. No.: 1,26e-18 Length: 158  
Score: 243.00 Matches: 56  
Percent Similarity: 56.49% Conservative: 31  
Best Local Similarity: 36.36% Mismatches: 55  
Query Match: 19.21% Indels: 12  
DB: 4 Gaps: 6

US-10-728-051-2 (1-717) x US-09-531-727-2 (1-158)

QY 5 ACCATCTAGTCCCTCGCCCTTTCTCTCGTCCCGCCAGCATCTGCGAGGCAGCAG 64  
Db 5 ThrileuLeuLeuSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23  
QY 65 TGGGAACCTCAAGAGAGACAGAGATCCAGAGCCAGCTCGAGAGCGGCGAACCTGAGGCC 124  
Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42  
QY 125 TGGAGCAACATCTATCGCAGAGATCCAA-----CGTGACGAGGATTCATATGAACGG 178  
Db 43 CysGluYshisIleMetGluYsIleGlnGlyArgGlyAspAspAspAspAsp 62  
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCT 238  
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrlleArgAsnGlu 82  
QY 239 GGATCTCTCAG-----CACAAAGAGAGGTGTTCATATGAGCTGAAC 280  
Db 83 GlyLysAspGluAspGluGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102  
QY 281 GAGTTTGAGAACACCAAGGTGCATGTGCGAGCATTCGACAGATTCAGGAGACCG 340

Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121  
QY 341 AGCATAGGTTCAGGAGGAGCAACAGGAGCAACAGATTCAAGAGGAGCTCAGGAACCTTG 400  
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysMetGluLysGluLeuIleAsnLeu 140  
QY 401 CCTCAACAGTGGCGCTTAGGCGACACACAGGCTTCGAGACTTG 442  
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154  
RESULT 8  
US-08-618-911-2  
Sequence 2, Application US/08618911  
GENERAL INFORMATION:  
APPLICANT: Jung, Rudolf  
APPLICANT: Hastings, Craig  
APPLICANT: Coughlan, Sean  
APPLICANT: Hu, David  
TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN SEEDS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.  
STREET: 700 Capital Square, 400 Locust Street  
CITY: Des Moines  
STATE: Iowa  
COUNTRY: USA  
ZIP: 50309  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/618,911  
FILING DATE: Concurrently herewith  
ATTORNEY/AGENT INFORMATION:  
NAME: Simon, Soma  
REGISTRATION NUMBER: 37,444  
REFERENCE/DOCKET NUMBER: 365-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (515) 248-4800  
TELEFAX: (515) 248-4844  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 155 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-618-911-2

Alignment Scores:  
Pred. No.: 1,85e-18 Length: 155  
Score: 241.50 Matches: 58  
Percent Similarity: 57.14% Conservative: 30  
Best Local Similarity: 37.66% Mismatches: 53  
Query Match: 19.09% Indels: 13  
DB: 2 Gaps: 7

US-10-728-051-2 (1-717) x US-08-618-911-2 (1-155)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGTCCCGCCAGCATCTGCGAGGCAG 61  
Db 4 LeuThrIleLeuLeuAlaLeuPheIle-----AlaHisThrCysCysAlaSer 21  
QY 62 CAGTGGGAACCTCAAGGACACAGAGATCCAGAGCCAGCTCGAGAGCGGCGAACCTGAGG 121  
Db 22 LysTrpGlnGlnHisGlnGlnGlnSerCysArgGluGlnLeuLysGlyIleAsnLeuAsn 41  
QY 122 CCTCGGACCAACATCTCATGCGAAGATCCAA-----CGTGACGAGGATTCATAT 172  
Db 42 ProCysGlu---HisIleMetGluLysIleGlnAlaGlyArgArgGlyGluAspGlySer 60



REGISTRATION NUMBER: 32,604  
REFERENCE/DOCKET NUMBER: ARK00895601B  
TELEPHONE: (501) 582-9111  
TELEFAX: (501) 521-4931  
TELEX: No. 6835824 applicable  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
STRANDEDNESS: No. 6835824 applicable  
TOPOLOGY: unknown  
MOLECULE TYPE: glycoprotein  
DESCRIPTION: identified as derived N-terminal  
HYPOTHETICAL: No  
ANTI-SENSE: No. 6835824 applicable  
FRAGMENT TYPE: N-terminal fragment  
ORIGINAL SOURCE:  
ORGANISM: Arachis hypogaea  
STRAIN: Florunner  
INDIVIDUAL ISOLATE: P38  
DEVELOPMENTAL STAGE:  
HABLOTYPE: No. 6835824 applicable  
TISSUE TYPE:  
CELL TYPE: No. 6835824 applicable  
CELL LINE: No. 6835824 applicable  
ORGANELLE: No. 6835824 applicable  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE: P38  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: No. 6835824 applicable  
MAP POSITION: No. 6835824 applicable  
UNITS: No. 6835824 applicable  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD: By agreement with  
IDENTIFICATION METHOD: protein information and established  
IDENTIFICATION METHOD: consensus sequence  
OTHER INFORMATION: Seed storage protein and  
OTHER INFORMATION: allergen  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-09-191-593-53

Alignment Scores:  
Pred. No.: 3 78e-09 Length: 28  
Score: 156.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.33% Indels: 0  
DB: 4 Gaps: 0

US-10-728-051-2 (1-717) x US-09-191-593-53 (1-28)

QY 59 CAGCAGTGGGAATCCCAAGAGACAGAGATGCCAGAGCGAGTCCGAGAGGGCGAACCTG 118  
Db 1 GlnGlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeu 20  
QY 119 AGCCCTGCGGACACATCTCATG 142

Db 21 ArgProCysGluGlnHisLeuMet 28  
RESULT 11  
US-08-670-186-4  
Sequence 4, Application US/08670186  
Patent No. 585343  
GENERAL INFORMATION:  
APPLICANT: SUN, SAMUEL S.M.  
APPLICANT: XIONG, LIWEN  
APPLICANT: HU, ZHONG  
APPLICANT: CHEN, HANG  
TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1888  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/670,186  
FILING DATE: 21-JUN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 23461-20007.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 822-0168  
TELEX: 90-4030 MRSNFOERSWSH  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-670-186-4

Alignment Scores:  
Pred. No.: 7 17e-09 Length: 158  
Score: 156.00 Matches: 46  
Percent Similarity: 47.56% Conservative: 32  
Best Local Similarity: 28.05% Mismatches: 52  
Query Match: 12.33% Indels: 34  
DB: 2 Gaps: 8

US-10-728-051-2 (1-717) x US-08-670-186-4 (1-158)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTTCCTCGTCCGCCAGCATCTGCG---AGG 59  
Db 5 IleLeuLeuLeuThrThrLeuAlaLeuPheValLeuLeuAlaAsnAlaSerIleTyArg 24  
QY 59 CAGCAGTGGGAATC-----CAAGGAGACAGAGATGCCAGAGC 97  
Db 25 ThrThrValGluLeuAspGluAspGlnProLeuCysArgArg 44  
QY 98 CAGCTCGAGAGG---CCGAACCTCGAGCAACATCTCATGCGAAGATCCAA 154  
Db 45 GlnPheGlnGlnHisLeuArgAlaCysGlnArgTyIleArgArgArgAlaGln 64  
QY 155 CQTGAC-----GAGGATTCATATGAACGGGACCCCTACAGCCCTAGTCAGATCCGTAC 208  
Db 65 ArgGlyGlyLeuValAspGluLeuGluAspValGluGluAsnGluAsp----- 82  
QY 209 AGCCCTAGTCCATATGATCGGAGCGCTGATCTCTCAGCACCAAGAGG----- 262







RESULT 15

US-09-191-593-58  
; Sequence 58, Application US/09191593  
; Patent No. 6835824  
; GENERAL INFORMATION:  
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,  
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,  
; APPLICANT: BANNON, Gary A  
; TITLE OF INVENTION: PEANUT ALLERGENS AND  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Head, Johnson & Kachigian  
; STREET: 112 W. Center St., Suite 230  
; CITY: Fayetteville  
; STATE: Arkansas AR  
; COUNTRY: United States of America  
; ZIP: 72701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: MS-DOS 5.2  
; SOFTWARE: Wordperfect 6.0C  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/191,593  
; FILING DATE: 13 NOVEMBER 1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/717,933  
; FILING DATE: 23 SEPTEMBER 1996  
; APPLICATION NUMBER: US 07/998,377  
; FILING DATE: 30 DECEMBER 1992  
; APPLICATION NUMBER: US 08/158,704  
; FILING DATE: 29 NOVEMBER 1993  
; APPLICATION NUMBER: US 60/009,455  
; FILING DATE: 29 DECEMBER 1995  
; APPLICATION NUMBER: US 08/610,424  
; FILING DATE: 04 MARCH 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ALEXANDER, DANIEL R  
; REGISTRATION NUMBER: 32,604  
; REFERENCE/DOCKET NUMBER: ARK00895601B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (501) 582-9111  
; TELEFAX: (501) 521-4931  
; TELEX: No. 6835824 applicable  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: No. 6835824 applicable  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; DESCRIPTION: identified as 17.5 kD N-terminal  
; DESCRIPTION: sequence of Ara h 2 p38 (gene sequence 20..47).  
; HYPOTHETICAL: No  
; ANTI-SENSE: No. 6835824 applicable  
; FRAGMENT TYPE: amino terminus  
; ORIGINAL SOURCE:  
; ORGANISM: Arachis hypogaea  
; STRAIN: Florunner  
; INDIVIDUAL ISOLATE: P38  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE: No. 6835824 applicable  
; TISSUE TYPE:  
; CELL TYPE: No. 6835824 applicable  
; CELL LINE: No. 6835824 applicable  
; ORGANELLE: No. 6835824 applicable  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE: P38

; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT: No. 6835824 applicable  
; MAP POSITION: No. 6835824 applicable  
; UNITS: No. 6835824 applicable  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD: By agreement with  
; IDENTIFICATION METHOD: protein information and established  
; IDENTIFICATION METHOD: consensus sequence  
; OTHER INFORMATION: Seed storage protein and  
; OTHER INFORMATION: allergen  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; TITLE:  
; JOURNAL:  
; VOLUME:  
; ISSUE:  
; PAGES:  
; DATE:  
; DOCUMENT NUMBER:  
; FILING DATE:  
; PUBLICATION DATE:  
; RELEVANT RESIDUES IN SEQ ID NO:  
US-09-191-593-58

Alignment Scores:  
Pred. No.: 3.95e-07 Length: 28  
Score: 138.00 Matches: 26  
Percent Similarity: 92.86% Conservative: 0  
Best Local Similarity: 92.86% Mismatches: 2  
Query Match: 10.91% Indels: 0  
DB: 4 Gaps: 0

US-10-728-051-2 (1-717) x US-09-191-593-58 (1-28)

QY	59	CAGCAGTGGAACTCCAAAGGACAGACAGATGCCAGAGCGAGCTCGAGAGGGCGACCTG	118
Db	1	GlnGlnTrpGluLeuGlnGlyAspArgArgGlnSerGlnLeuGluArgAlaAsnLeu	20
QY	119	AGGCCCTGCGAGCAACATCTCATG	142
Db	21	ArgProCysGluGlnHisLysMet	28

Search completed: April 8, 2005, 05:34:11  
Job time : 22.3806 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 8, 2005, 05:25:59 ; Search time 82.3048 Seconds  
(without alignments)  
5784.401 Million cell updates/sec

Title: US-10-728-051-2  
Perfect score: 1265  
Sequence: 1 gccaccatactagtagcc.....taaaagatcatgttttgtt 717

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 2836020

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool.p/US10728051/runat\_07042005\_125048\_14128/app\_query.fasta\_1.4757  
-DB=Published Applications AA -QFMT=fastn -SURFIX=rapb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10728051@cgn2\_1\_663 @runat\_07042005\_125048\_14128  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEOURRY -NEG SCORES=0 -WAIT -DGBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

Alignment Scores: 9.72e-80 Length: 157  
Pred. No.: 842.00 Matches: 157  
Score:

No.	Score	Match	Length	DB	ID	Description
1	842	66.6	157	14	US-10-228-806-4	Sequence 4, Appl
2	842	66.6	157	15	US-10-100-303A-63	Sequence 63, Appl
3	835	66.0	160	17	US-10-899-551-4	Sequence 4, Appl
4	819	64.7	156	15	US-10-245-871-53	Sequence 53, Appl
5	819	64.7	156	15	US-10-253-286-53	Sequence 53, Appl
6	771	60.9	166	15	US-10-100-303A-82	Sequence 82, Appl
7	760	60.1	167	17	US-10-899-551-56	Sequence 56, Appl
8	699	55.3	166	9	US-09-731-221-77	Sequence 77, Appl
9	463	36.6	83	15	US-10-100-303A-81	Sequence 81, Appl
10	243	19.2	158	15	US-10-302-633-2	Sequence 2, Appl
11	243	19.2	158	15	US-10-424-599-169225	Sequence 169225,
12	241.5	19.1	155	15	US-10-424-599-284372	Sequence 284372,
13	177.5	14.0	168	14	US-10-165-289A-2	Sequence 2, Appl
14	149.5	11.8	167	16	US-10-382-066-29	Sequence 29, Appl
15	144	11.4	169	14	US-10-165-289A-4	Sequence 4, Appl
16	139	11.0	184	11	US-09-782-130-3	Sequence 3, Appl
17	139	11.0	184	14	US-10-100-121-48	Sequence 48, Appl
18	129.5	10.2	150	15	US-10-442-174A-1	Sequence 1, Appl
19	127.5	10.1	151	15	US-10-424-599-167336	Sequence 167336,
20	127	10.0	22	14	US-10-228-806-67	Sequence 67, Appl
21	126	10.0	102	15	US-10-424-599-219338	Sequence 219338,
22	119	9.4	20	9	US-09-731-221-66	Sequence 66, Appl
23	118	9.2	261	15	US-10-291-265-715	Sequence 715, App
24	117	9.2	20	9	US-09-731-221-63	Sequence 63, Appl
25	116	9.2	20	9	US-09-731-221-60	Sequence 60, Appl
26	115	9.1	20	9	US-09-731-221-61	Sequence 61, Appl
27	113	8.9	20	9	US-09-731-221-65	Sequence 65, Appl
28	112.5	8.9	101	15	US-10-424-599-275916	Sequence 275916,
29	112	8.9	20	9	US-09-731-221-54	Sequence 54, Appl
30	112	8.9	20	9	US-09-731-221-59	Sequence 59, Appl
31	112	8.9	20	9	US-09-731-221-62	Sequence 62, Appl
32	112	8.9	20	9	US-09-731-221-64	Sequence 64, Appl
33	111	8.8	20	9	US-09-731-221-57	Sequence 57, Appl
34	111	8.8	20	9	US-09-731-221-74	Sequence 74, Appl
35	111	8.8	20	9	US-09-731-221-75	Sequence 75, Appl
36	111	8.8	20	9	US-09-731-221-76	Sequence 76, Appl
37	110	8.7	25	15	US-10-245-871-60	Sequence 60, Appl
38	110	8.7	25	15	US-10-253-286-60	Sequence 60, Appl
39	109	8.6	20	15	US-10-245-871-55	Sequence 55, Appl
40	109	8.6	20	15	US-10-253-286-55	Sequence 55, Appl
41	108	8.5	20	9	US-09-731-221-53	Sequence 53, Appl
42	108	8.5	20	9	US-09-731-221-56	Sequence 56, Appl
43	108	8.5	20	9	US-09-731-221-67	Sequence 67, Appl
44	108	8.5	20	9	US-09-731-221-68	Sequence 68, Appl
45	108	8.5	50	15	US-10-424-599-250084	Sequence 250084,

ALIGNMENTS

RESULT 1  
US-10-228-806-4  
; Sequence 4, Application US/10228806  
; Publication No. US20030049237A1  
; GENERAL INFORMATION:  
; APPLICANT: Bannan, et al.  
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions  
; TITLE OF INVENTION: to Allergy  
; FILE REFERENCE: 2002834-0043  
; CURRENT APPLICATION NUMBER: US/10/228, 806  
; CURRENT FILING DATE: 2002-08-26  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 157  
; TYPE: PRT  
; ORGANISM: Arachis hypogaea  
US-10-228-806-4

Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 66.56%  
DB: 14  
Indels: 0  
Gaps: 0  
Conservative: 0

US-10-728-051-2 (1-717) x US-10-228-806-4 (1-157)

QY 2 CTCACCATAGTACGCTCGCCCTTTTCCTCCTCGTCCCAACGATCTGCGAGGAG 61  
DB 1 LeuThriLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaAArgGln 20  
QY 62 CAGTGGGAACCTCAAGGAGACAGAAGATGCCAGGCGAGCTCGAGAGGCGCACTGAGG 121  
DB 21 GlnTrpGluLeuGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArg 40  
QY 122 CCCTGGAGCAACATCTCATGACAGATCCCAACGTGACGAGATTCATATGAACGGGAC 181  
DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArg 60  
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGGCGCTGGA 241  
DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80  
QY 242 TCCTCTCAGCACCAAGAGGCTGTCGAATGAGCTGAACGAGTTTGAGAACACCAAGG 301  
DB 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 100  
QY 302 TGCATGTGCGAGGCATTCGAACAGATCATGAGACCATGAGGATAGTTGACAGGGGAGG 361  
DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgGlyAlaGly 120  
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTTCAGAACTTTCCTCAACAGTGCAGGCGCTTAGG 421  
DB 121 GlnGlnGluGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140  
QY 422 GCACCAAGGCTTGCAGCTTGCAGCTCGAAAGTGGCGGCGAGACAGATAC 472  
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

## RESULT 2

US-10-100-303A-63

; Sequence 63, Application US/10100303A  
; Publication No. US20030202980A1  
; GENERAL INFORMATION:  
; APPLICANT: Caplan, et al.  
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction  
; FILE REFERENCE: 2002834-0166  
; CURRENT FILING DATE: US/10/100,303A  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 63  
; LENGTH: 157  
; TYPE: PRT  
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2  
US-10-100-303A-63

## Alignment Scores:

Pred. No.: 9,72e-80 Length: 157  
Score: 842.00 Matches: 157  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 66.56%  
DB: 15  
Indels: 0  
Gaps: 0  
Conservative: 0

US-10-728-051-2 (1-717) x US-10-100-303A-63 (1-157)

QY 2 CTCACCATAGTACGCTCGCCCTTTTCCTCCTCGTCCCAACGATCTGCGAGGAG 61  
DB 1 LeuThriLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaAArgGln 20  
QY 62 CAGTGGGAACCTCAAGGAGACAGAAGATGCCAGGCGAGCTCGAGAGGCGCACTGAGG 121

DB 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40  
QY 122 CCCTGGAGCAACATCTCATGACAGATCCCAACGTGACGAGATTCATATGAACGGGAC 181  
DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArg 60  
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGGCGCTGGA 241  
DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80  
QY 242 TCCTCTCAGCACCAAGAGGCTGTCGAATGAGCTGAACGAGTTTGAGAACACCAAGG 301  
DB 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 100  
QY 302 TGCATGTGCGAGGCATTCGAACAGATCATGAGACCATGAGGATAGTTGACAGGGGAGG 361  
DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgGlyAlaGly 120  
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTTCAGAACTTTCCTCAACAGTGCAGGCGCTTAGG 421  
DB 121 GlnGlnGluGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140  
QY 422 GCACCAAGGCTTGCAGCTTGCAGCTCGAAAGTGGCGGCGAGACAGATAC 472  
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

## RESULT 3

US-10-899-551-4

; Sequence 4, Application US/10899551  
; Publication No. US20050063994A1  
; GENERAL INFORMATION:  
; APPLICANT: Burks, A. Wesley  
; APPLICANT: Caplan, Michael J.  
; APPLICANT: Sampson, Hugh A.  
; APPLICANT: Howard, Soshin B.  
; APPLICANT: Bottomly, Kim H.  
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy  
; FILE REFERENCE: 2002834-0233  
; CURRENT FILING DATE: US/10/899,551  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 160  
; TYPE: PRT  
; ORGANISM: species Arachis hypogaea  
US-10-899-551-4

## Alignment Scores:

Pred. No.: 5.35e-79 Length: 160  
Score: 835.00 Matches: 156  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 66.01%  
DB: 17  
Indels: 0  
Gaps: 0  
Conservative: 0

US-10-728-051-2 (1-717) x US-10-899-551-4 (1-160)

QY 2 CTCACCATAGTACGCTCGCCCTTTTCCTCCTCGTCCCAACGATCTGCGAGGAG 61  
DB 4 LeuThriLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaAArgGln 23  
QY 62 CAGTGGGAACCTCAAGGAGACAGAAGATGCCAGGCGAGCTCGAGAGGCGCACTGAGG 121  
DB 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43  
QY 122 CCCTGGAGCAACATCTCATGACAGATCCCAACGTGACGAGATTCATATGAACGGGAC 181  
DB 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArg 63  
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGGCGCTGGA 241



```
Db      64  ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly  83
Qy      242  TCCTCTCAGCAACCAAGAGAGGTGTGCAATGAGCTGAACAGTTTGAGAAACAACCAAGG  301
Db      84  SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg  103
Qy      302  TGCATGTGCCAGGACCATTCGCAACAGATCATCGAGAACCCAGAGCGATAGTTGCCAGGGGAGG  361
Db      104  CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg  123
Qy      362  CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG  421
Db      124  GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg  143
Qy      422  GCACCAAGGCTTGCAGCTTGGACGTGCGGAGGCGGAGACAGA  469
Db      144  AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArg  159

RESULT 4
US-10-245-871-53
; Sequence 53, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-245-871-53

Alignment Scores:
Pred. No.:      2,59e-77      Length:      156
Score:          819.00      Matches:      153
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    64.74%      Indels:      0
DB:             15      Gaps:        0

US-10-728-051-2 (1-717) x US-10-245-871-53 (1-156)
Qy      2  CTCACCACTAGTAGCCCTCGCCCTTTTCTCTCTCGCTGCCACGCGATCTGGAGGCGAG  61
Db      4  LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln  23
Qy      62  CAGTGGGAATCCCAAGGAGACAGAAGATGCCAGAGCGAGCTCGAGAGGGCGAACCCTGAGG  121
Db      24  GlnTrpGluLeuGlnArgAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg  43
Qy      122  CCCTGCGAGCAACATCTCATGCAAGAGATCCAACTGACGAGGATTCATATGAACGGGAC  181
Db      44  ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp  63
Qy      182  CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGGCGCTGGA  241
Db      64  ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly  83
Qy      242  TCCTCTCAGCAACCAAGAGAGGTGTGCAATGAGCTGAACAGTTTGAGAAACAACCAAGG  301
Db      84  SerSerGlnHisGlnGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp  63
Qy      182  CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGGCGCTGGA  241
Db      64  ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly  83
Qy      242  TCCTCTCAGCAACCAAGAGAGGTGTGCAATGAGCTGAACAGTTTGAGAAACAACCAAGG  301
Db      84  SerSerGlnHisGlnGlnArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg  103
Qy      302  TGCATGTGCCAGGACCATTCGCAACAGATCATCGAGAACCCAGAGCGATAGTTGCCAGGGGAGG  361
```

```
Db      104  CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg  123
Qy      362  CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG  421
Db      124  GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg  143
Qy      422  GCACCAAGGCTTGCAGCTTGGACGTGCGGAGGCGGAGGCGGC  460
Db      144  AlaProGlnArgCysAspLeuAspValGluSerGlyGly  156

RESULT 5
US-10-253-286-53
; Sequence 53, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-253-286-53

Alignment Scores:
Pred. No.:      2,59e-77      Length:      156
Score:          819.00      Matches:      153
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    64.74%      Indels:      0
DB:             15      Gaps:        0

US-10-728-051-2 (1-717) x US-10-253-286-53 (1-156)
Qy      2  CTCACCACTAGTAGCCCTCGCCCTTTTCTCTCTCGCTGCCACGCGATCTGGAGGCGAG  61
Db      4  LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln  23
Qy      62  CAGTGGGAATCCCAAGGAGACAGAAGATGCCAGAGCGAGCTCGAGAGGGCGAACCCTGAGG  121
Db      24  GlnTrpGluLeuGlnArgAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg  43
Qy      122  CCCTGCGAGCAACATCTCATGCAAGAGATCCAACTGACGAGGATTCATATGAACGGGAC  181
Db      44  ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp  63
Qy      182  CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGGCGCTGGA  241
Db      64  ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly  83
Qy      242  TCCTCTCAGCAACCAAGAGAGGTGTGCAATGAGCTGAACAGTTTGAGAAACAACCAAGG  301
Db      84  SerSerGlnHisGlnGlnArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg  103
Qy      302  TGCATGTGCCAGGACCATTCGCAACAGATCATCGAGAACCCAGAGCGATAGTTTGAGAGGGGAGG  361
Db      104  CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg  123
Qy      362  CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG  421
Db      124  GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg  143
Qy      422  GCACCAAGGCTTGCAGCTTGGACGTGCGGAGGCGGAGGCGGC  460
```

Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156

## RESULT 6

US-10-100-303A-82  
; Sequence 82, Application US/10100303A  
; Publication No. US20030202980A1

GENERAL INFORMATION:  
; APPLICANT: Caplan, et al.

; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction  
; TITLE OF INVENTION: to Allergy

; FILE REFERENCE: 2002834-0166

; CURRENT APPLICATION NUMBER: US/10/100,303A

; CURRENT FILING DATE: 2002-03-16

; NUMBER OF SEQ ID NOS: 138

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 82

; LENGTH: 166

; TYPE: PRT

; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2

US-10-100-303A-82

## Alignment Scores:

Pred. No.:	3.06e-72	Length:	166
Score:	771.00	Matches:	141
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	60.95%	Indels:	0
DB:	15	Gaps:	0

US-10-728-051-2 (1-717) x US-10-100-303A-82 (1-166)

QY	50	TCCTCGAGGCGAGCTGGGAACCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGG	109
Db	15	SerAlaArgGlnGlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArg	34
QY	110	GCGAACCTGAGGCGCTCGAGCAACATCTCATGCAGAAATCAACGTGACGAGGATTCA	169
Db	35	AlaAsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSer	54

QY	170	TATGAACGGAGCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGG	229
Db	55	TyrGluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg	74
QY	230	AGAGGCGCTGATCCTCTCAGCACCAGAGAGGTGTTCATAGCTGAACGAGTTTGA	289
Db	75	ArgGlyAlaGlySerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGlu	94

QY	290	AACAACCAAGGTGCATGTGCGAGGCATTGCAACAGATCATGCAGAACACAGACGATAGG	349
Db	95	AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg	114

QY	350	TTGAGGGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAG	409
Db	115	LeuGlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGln	134

QY	410	TGCGGCTTAGGACACACAGCCTTGCAGCTTGAGCTGGAAGTGGCGGAGACAGACA	469
Db	135	CysGlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg	154

QY 470 TAC 472

Db 155 Tyr 155

## RESULT 7

US-10-899-551-56

; Sequence 56, Application US/10899551

; Publication No. US20050063994A1

GENERAL INFORMATION:

; APPLICANT: Caplan, Michael J.

; APPLICANT: Burke, A. Wesley

; APPLICANT: Sampson, Hugh B.

; APPLICANT: Howard, Sosin B.

; APPLICANT: Bottomly, Kim H.

; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy  
; FILE REFERENCE: 2002834-0233

; CURRENT APPLICATION NUMBER: US/10/899,551

; CURRENT FILING DATE: 2004-07-26

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 56

; LENGTH: 167

; TYPE: PRT

; ORGANISM: species Arachis hypogaea

US-10-899-551-56

## Alignment Scores:

Pred. No.:	4.43e-71	Length:	167
Score:	760.00	Matches:	139
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	60.08%	Indels:	0
DB:	17	Gaps:	0

US-10-728-051-2 (1-717) x US-10-899-551-56 (1-167)

QY	53	GGAGGCGAGCTGGGAACCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCG	112
Db	17	AlaArgGlnGlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAla	36
QY	113	AACCTGAGGCGCTCGAGCAACATCTCATGCAGAAATCAACGTGACGAGGATTCAAT	172
Db	37	AsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyr	56

QY	173	GAACGGAGCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGA	232
Db	57	GluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArg	76
QY	233	GGCGCTGATCCTCTCAGCACCAGAGAGGTGTTCATGCAATGAGCTGAACGAGTTTGA	292
Db	77	GlyAlaGlySerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsn	96

QY	293	AACCAAGGTGCATGTGCGAGGCATTGCAACAGATCATGCAGAACACAGAGCCATAGGTTG	352
Db	97	AsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeu	116
QY	353	CAGGGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCCAAACAGTGC	412
Db	117	GlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCys	136

QY	413	GGCCTTAGGCGCACACAGCGTTGCGACTTGAGCGTCGAAAGTGGCGGCGAGACACAGA	469
Db	137	GlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg	155

## RESULT 8

US-09-731-221-77

; Sequence 77, Application US/09731221

; Patent No. US20020018778A1

GENERAL INFORMATION:

; APPLICANT: Caplan, Michael

; TITLE OF INVENTION: Passive Desensitization

; FILE REFERENCE: 2002834-0103

; CURRENT APPLICATION NUMBER: US/09/731,221

; CURRENT FILING DATE: 2001-12-06

; NUMBER OF SEQ ID NOS: 79

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 77

; LENGTH: 166

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:Arachis

US-09-731-221-77

## Alignment Scores:

Pred. No.:	1.21e-64	Length:	166
------------	----------	---------	-----

Score: 699.00 Matches: 131  
Percent Similarity: 92.91% Conservatives: 0  
Best Local Similarity: 92.91% Mismatches: 10  
Query Match: 55.26% Indels: 0  
DB: 9 Gaps: 0

US-10-728-051-2 (1-717) x US-09-731-221-77 (1-166)

QY 50 TCTGCGAGGAGGAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGG 109  
Db 15 SerAlaArgGlnGlnAlaGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuAlaArg 34  
QY 110 GCGAAGCTGAGGCGCTCGGAGCAACATCTCATCAGAGATCCAACTGACGAGGATTCA 169  
Db 35 AlaAsnLeuArgAlaCysGluAlaHisLeuMetGlnLysIleGlnAlaAspGluAspSer 54  
QY 170 TATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGG 229  
Db 55 TyrGluArgAlaProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArg 74  
QY 230 AGAGGCGCTGGATCCTCTCAGCAACCAAGAGAGGTGTTCATAGCTGAAACGAGTTTGAG 289  
Db 75 ArgGlyAlaGlySerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGlu 94  
QY 290 AACAAACCAAGGTGCATGTGCGAGGCATTCACAGATCATGAGACCAAGCGGATAGG 349  
Db 95 AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg 114  
QY 350 TTGCGGGGAGGCAACAGGAGCAACAGTTCACAGAGGAGCTCAGGAACCTTCGCTCAACAG 409  
Db 115 LeuGlnGlyAlaGlnGlnGlnGlnPheGlyArgGluAlaArgAsnLeuProGlnGln 134  
QY 410 TCGGCGCTTAGGCACACAGCGTTGCGACTTGGAAGTCGAAAGTGGCGGCAGAGACAGA 469  
Db 135 CysGlyLeuArgAlaProGlnArgCysAspAlaAspValGluSerGlyGlyArgAspArg 154  
QY 470 TAC 472  
Db 155 Tyr 155

RESULT 9  
US-10-100-303A-81  
; Sequence 81, Application US/10100303A  
; Publication No. US20030202980A1  
; GENERAL INFORMATION:  
; APPLICANT: Caplan, et al.  
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction  
; FILE REFERENCE: 2002834-0166  
; CURRENT APPLICATION NUMBER: US/10/100,303A  
; CURRENT FILING DATE: 2002-03-18  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 81  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2  
US-10-100-303A-81

Alignment Scores:  
Pred. No.: 7.94e-40 Length: 83  
Score: 463.00 Matches: 83  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 36.60% Indels: 0  
DB: 15 Gaps: 0

US-10-728-051-2 (1-717) x US-10-100-303A-81 (1-83)

QY 68 GAATCCCAAGGAGACAGAGATGCCAGGCTCGAGAGGGCGCAACCTGAGGCCCTGC 127  
Db 1 GluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArgProCys 20

QY 128 GAGCACATCTCATGCGAGATCCAACTGAGAGGATTCATATGACGGGACCCGTAC 187  
Db 21 GluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAspProTyr 40  
QY 188 AGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGATCCTCT 247  
Db 41 SerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGlySerSer 60  
QY 248 CAGCACCAAGAGAGGTGTTCATAGCTGAAACGAGTTTGAGAACCAACCAAGGTTGATG 307  
Db 61 GlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArgCysMet 80  
QY 308 TCGAGGCA 316  
Db 81 CysGluAla 83

RESULT 10  
US-10-302-633-2  
; Sequence 2, Application US/10302633  
; Publication No. US20030229038A1  
; GENERAL INFORMATION:  
; APPLICANT: de Lumen, Benito O.  
; Galvez, Alfredo F.  
; TITLE OF INVENTION: Luminasin Peptides  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 75 DENISE DRIVE  
; CITY: HILLSBOROUGH  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94010  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/302,633  
; FILING DATE: 22-No. US20030229038A1-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/531,727  
; FILING DATE: 21-Mar-2000  
; APPLICATION NUMBER: 08/938,675  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B98-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 343-4341  
; TELEFAX: (650) 343-4342  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 158 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-302-633-2

Alignment Scores:  
Pred. No.: 1.54e-16 Length: 158  
Score: 243.00 Matches: 56  
Percent Similarity: 56.49% Conservatives: 31  
Best Local Similarity: 36.36% Mismatches: 55  
Query Match: 19.21% Indels: 12  
DB: 15 Gaps: 6

US-10-728-051-2 (1-717) x US-10-302-633-2 (1-158)



```
QY 233 GCGCTGATCCTCTCAG-----CACCAAGAGAGGTGTTGCAATGAGCTGAAC 280
Db 80 LysGluGlyLysGluGluGluGluGlyHisMetGlnLysCysSerGluMetSer 99
QY 281 GAGTTTGAGAACCAACAAAGTGTGATGCGAGGATTGCAACAGATCATGGAGAACAG 340
Db 100 GluLeuLys---SerProIleCysGlnCysAlaLeuGlnLysIleMetAspAsnGln 118
QY 341 AGCATAGTTGTCAGGGAGGCAACAGAGCAACAGTTTCAAGAGGGAGCTCAGAACTTG 400
Db 119 SerGluGlnLeuGluGlyLys--GluLysLysGlnMetGluArgGluLeuMetAsnLeu 137
QY 401 CCTCAACAGTGGCGCTTAGGGCACACAGCGTTGCGACTTG 442
Db 138 AlaIleArgCysArgLeuGlyProMetIleGlyCysAspLeu 151

RESULT 13
US-10-165-289A-2
; Sequence 2, Application US/10165289A
; Publication No. US20030159174A1
; GENERAL INFORMATION:
; APPLICANT: QIU, Xiao
; APPLICANT: TRUSKA, Martin
; APPLICANT: HU, Zhiyuan
; TITLE OF INVENTION: Flax (Linum usitatissimum L.)
; FILE REFERENCE: Seed-Specific Promoters
; CURRENT FILING DATE: 2002-06-06
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Linum usitatissimum
US-10-165-289A-2

Alignment Scores:
Pred. No.: 1,27e-09 Length: 168
Score: 177.50 Matches: 54
Percent Similarity: 45.71% Conservative: 26
Best Local Similarity: 30.86% Mismatches: 48
Query Match: 14.03% Indels: 47
DB: 14 Gaps: 9

US-10-728-051-2 (1-717) x US-10-165-289A-2 (1-168)
QY 2 CTCACCATCTAGTAGCCCTCGCC-----CTTTCTCTCTCGCTGCCCGCATCT 52
Db 4 LeuMetSerLeuAlaValAlaThrAlaPheLeuPheLeuIleValValAspAlaSer 23
QY 53 GCGAGGAGCAGTGGGAATCCRAAGAGAC----- 82
Db 24 ValArgThrValIleIleAspGluAspThrAsnGlnGlyArgGlyGlnGly 43
QY 83 -----AGAGATGCCAGAGCAGCTC---GAGAGGGCGAACCTGAGGCCCTGC 127
Db 44 GlnGlyGlnGlnGlnCysGluLysGlnIleGlnGlnAspTyrLeuArgSerCys 63
QY 128 GAGCAACATCTCAGAGAGATCCACGTGACAGGATTTCATATGAACCGGACCCGTAC 187
Db 64 GlnGlnPheLeuTrpGluLysValGlnLysGlyArgSerTyr----- 78
QY 188 AGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGGCTGGA----- 241
Db 79 -----TyrTrAsnGlnGlyArgGlyGly 87
QY 242 ----TCCTCTCAGCACCAGAGAGGTGTTGCAATGAGCTGAACAGTGTGAGAACCAACAA 298
Db 88 GlnGlnSerGlnHisPheAspSerCysCysAspAspLeuLysGlnLeuArgSer----- 105
```

```
QY 299 AGGTGCATGTGCGAGGATTGCAACAGATCATG-----CAGAACCCAGAGCGATAGTGTG 352
Db 106 GluCysThrCysArgGlyLeuGluArgAlaIleGlyGlnMetArgGlnAspIleGlnGln 125
QY 353 CAGCGGAGGCAACAGGAG-----CAACAGTTCAAGAGGAGCTCAGGAAC 397
Db 126 GlnGlyGlnGlnGlnGlnValGluArgTrpValGlnGlnAlaLysGlnValAlaArgAsp 145
QY 398 TTCCCTCAACAGTGGCGCTTAGGGCACCAACAGCGTTGCGACTTG 442
Db 146 LeuProGlyGlnCysGlyThrGln--ProSerArgCysGlnLeu 159

RESULT 14
US-10-382-066-29
; Sequence 29, Application US/10382066
; Publication No. US20040177404A1
; GENERAL INFORMATION:
; APPLICANT: Li, ZhiJian
; APPLICANT: Gray, Dennis
; TITLE OF INVENTION: Nucleotide Sequences of 2S Albumin Gene and its Promoter from Grape
; FILE REFERENCE: Uses Thereof
; CURRENT FILING DATE: 2003-03-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Vitis vinifera cv. Merlot
US-10-382-066-29

Alignment Scores:
Pred. No.: 1,14e-06 Length: 167
Score: 149.50 Matches: 43
Percent Similarity: 44.51% Conservative: 34
Best Local Similarity: 24.86% Mismatches: 51
Query Match: 11.82% Indels: 45
DB: 16 Gaps: 7

US-10-728-051-2 (1-717) x US-10-382-066-29 (1-167)
QY 2 CTCACCATCTAGTAGCCCTCGCCCTTTCTCTCTCGCT---GCCACGCGATCTGCGAGG 58
Db 4 LeuSerIlePheAlaIleThrLeuLeuLeuLeuAlaIleSerAsnAlaThrIleTyr 23
QY 59 CAG-----CAGTGGGAATCCCAAGGA 79
Db 24 GlnThrThrValIleThrArgAspAspGlySerGluPheGlyGlnPheGlnGlySerGln 43
QY 80 GACAGAGATGCGAGAGCCAGCTCGAGAGGGCGAACCTGAGGCCCTGCGAGCAACATCTC 139
Db 44 SerGlnArgCysArgGlnGlnIleGlnGlnPheGlnGlnCysGluArgTyrIle 63
QY 140 ATCAGAGAATCCACGTGACGAGGATTCATATGAACGGGACCCGTACACCCCTAGTCAG 199
Db 64 ArgGlnGlnAlaGluGlnGlnGlnGlyGly----- 73
QY 200 GATCCGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTGCATCCTCTCAGACCAAGAG 259
Db 74 -----GlnGlyAspValLeuIleLeuArgGlyIleArgAsnGlnGlnGlnGlu 90
QY 260 -----AGGTGTTTCCAATGAGCTGAACGAGTTTGAGAACCAACCAAGGTGC 304
Db 91 GlnGlnTrpLeuArgGlnCysCysGlnAlaLeuGlnAsnMetAsp-----GlnGlnCys 108
QY 305 ATGTGGCGGCGATTCGAACAGATCATCGAGAACACAGAGCGATAGGTGCGAGGG----- 358
Db 109 GlnCysGluGlyLeuArgGlnIleValGlnArgGlnGlnGlnGlyGlnGlyGlnGly 128
QY 359 -----AGGCAACAGGAGCAACAGTTTCAAGAGGAGCTCAGGAAC 397
Db 129 GlnGlyGlnGlyGlnGlyGlnArgGluGlnGlnGlnGlnMetMetGlnIleAlaGlnLys 148
```

QY 398 TTGCTCAACAGTGGCGCTTAGGCACACACAGCGTTGC 436  
|||||:|||||  
Db 149 LeuProGluArgCysGly-----SerGlyGlnAlaCys 159  
|||:|||||

## RESULT 15

US-10-165-289A-4  
; Sequence 4, Application US/10165289A  
; Publication No. US20030159174A1  
; GENERAL INFORMATION:  
; APPLICANT: QIU, Xiao  
; APPLICANT: TRUSKA, Martin  
; APPLICANT: HU, Zhiyuan  
; TITLE OF INVENTION: Flax (Linum usitatissimum L.)  
; FILE REFERENCE: BNZ-005  
; CURRENT APPLICATION NUMBER: US/10/165,289A  
; CURRENT FILING DATE: 2002-06-06  
; PRIOR APPLICATION NUMBER: 60/295823  
; PRIOR FILING DATE: 2001-06-06  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 169  
; TYPE: PRT  
; ORGANISM: Linum usitatissimum  
US-10-165-289A-4

Alignment Scores:  
Pred. No.: 4.37e-06 Length: 169  
Score: 144.00 Matches: 48  
Percent Similarity: 45.40% Conservative: 31  
Best Local Similarity: 27.59% Mismatches: 47  
Query Match: 11.38% Indels: 48  
DB: 14 Gaps: 10

US-10-728-051-2 (1-717) x US-10-165-289A-4 (1-169)

QY 2 CTCACCTACTAGTAGCGCTCGCC-----CTTTTCCTCTCGCTGCCACGCACTCT 52  
Db 4 LeuMetSerLeuAlaAlaValAlaThrAlaPheLeuPheLeuValValAspAlaSer 23  
|||:|||||  
QY 53 GCGAGGCGACGAGTGGGAATC-----73  
|||:|||||  
Db 24 ValArgThrThrValIleIleAspGluGluThrAsnGlnGlyArgGlyGlyGlnGly 43  
|||:|||||  
QY 74 ---CAAGGAGACAGAGA---TGCCAGAGCCAGCTCGAGAGGCGAAC---CTGAGGCC 124  
|||||:|||||  
Db 44 GlyGlnGlyGlnGlnGlnSerCysGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 63  
|||||:|||||  
QY 125 TCGAGCAACATCTCATGCGAGAGATCCACGTGACGAGGATTCATGACGGGACCCG 184  
|||||:|||||  
Db 64 CysGlnGlnPheMetTrpGluLysValGlnArgGlyGlyArgSerHis-----79  
|||||:|||||  
QY 185 TACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCGAGAGCGCTGGA---241  
|||:|||||  
Db 80 -----TyrTrpAsnGlnGlyArgGlyGly 87  
|||:|||||  
QY 242 -----TCCTCTAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACAA 295  
|||||:|||||  
Db 88 GlyGluGlnSerGlnTyrPheAspSerCysCysAspLeuLysGlnLeuSerThrGly 107  
|||||:|||||  
QY 296 CAAAGGTGCATGTGCGAGGATTCACACAGATCATG-----CAGAACCGAGCGATAGG 349  
|||:|||||  
Db 108 -----CysThrCysArgGlyLeuGluuArgAlaIleGlyGlnMetArgGlnGluIleGln 125  
|||:|||||  
QY 350 TTCAGGGGAGGACACAGAG-----CAACAGTTCAAGAGGAGCTCAGG 394  
|||||:|||||  
Db 126 GlnGlnGlyGlnGlnGlnValGlnArgTrpIleGlnGlnAlaLysGlnIleAlaLys 145  
|||||:|||||  
QY 395 AACTGCTCAACAGTGGCGCTTAGGGACACACAGCGTTGC 436  
|||:|||||  
Db 146 AspLeuProGlyGlnCysArgThrGln---ProSerGlnCys 158  
|||:|||||

Search completed: April 8, 2005, 06:48:53  
Job time : 85.3048 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model.

Run on: April 8, 2005, 03:43:40 ; Search time 95.8964 Seconds  
(without alignments)

5783.472 Million cell updates/sec

Title: US-10-728-051-2

Perfect score: 1265

Sequence: 1 gctccaccatactagtagcc.....taaaagaatcatgttttgtt 717

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10728051/runat\_07042005\_125044\_14041/app.query.fasta\_1.4757  
-DB=A.Geneseq -QFWT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOPFCI=0.1 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10728051@cgn\_1.1.672 @runat\_07042005\_125044\_14041 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.Geneseq\_16Dec04.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	842	66.6	157	2 AAW24153	Aaw24153 Peanut al
2	842	66.6	157	2 AAW24164	Aaw24164 Peanut al
3	842	66.6	157	2 AAY15245	Aay15245 Peanut al
4	842	66.6	157	6 AAU04707	Aau04707 Anaphylac
5	842	66.6	157	6 ABUS2464	Abu52464 Peanut Ar
6	842	66.6	157	7 ADG27518	Adg27518 Peanut al
7	838.5	66.3	207	4 AAB82383	Aab82383 Peanut al
8	830	65.6	156	2 AAY40973	Aay40973 Ara h 2 p
9	819	64.7	156	8 ADO38314	Ado38314 Ara h 2 M
10	814	64.3	157	6 ABUS2576	Abu52576 Peanut Ar

11	812	64.2	156	8 ADM12096	Adm12096 Arachis h
12	812	64.2	157	6 ABUS2575	Abu52575 Peanut Ar
13	774	61.2	157	6 ABUS2577	Abu52577 Peanut Ar
14	771	60.9	166	6 AAY40968	Aay40968 Recombina
15	771	60.9	166	6 ABUS2482	Abu52482 Peanut Ar
16	771	60.9	166	7 ADG27536	Adg27536 T7/Hig-ta
17	699	55.3	166	3 AAB33600	Aab33600 Modified
18	699	55.3	166	4 AAU05035	Aau05035 Modified
19	688.5	54.4	167	4 AAU04710	Aau04710 Modified
20	243	19.2	158	2 AAW23419	Aaw23419 Soybean a
21	243	19.2	158	2 AAY05723	Aay05723 Soybean G
22	243	19.2	158	7 ADH89269	Adh89269 G.max 2S
23	243	19.2	158	8 ADG44004	Adg44004 G.max 2S
24	241.5	19.1	155	2 AAW23418	Aaw23418 Soybean a
25	241.5	19.1	155	7 ADH89267	Adh89267 G.max na
26	241.5	19.1	155	8 ADG44002	Adg44002 G.max 2S
27	229	18.1	158	2 AAW23420	Aaw23420 Chimeric
28	184.5	14.6	323	7 ADH89277	Adh89277 Sunflower
29	184.5	14.6	323	8 ADG44012	Adg44012 H. annuus
30	177.5	14.0	168	6 AAE36076	Aae36076 Flax Conl
31	156	12.3	28	2 AAW24194	Aaw24194 Peanut al
32	156	12.3	158	2 AAW23586	Aaw23586 Mabinlin
33	151	11.9	158	2 AAW23588	Aaw23588 Mabinlin
34	151	11.9	295	7 ADH89275	Adh89275 Sunflower
35	151	11.9	295	8 ADG44010	Adg44010 H. annuus
36	149.5	11.8	167	8 ADS16393	Ads16393 Grape 2S
37	146.5	11.6	164	1 AAP91892	Aap91892 1kb fragm
38	146.5	11.6	164	1 AAP96144	Aap96144 Sequence
39	146.5	11.6	164	7 ADH89227	Adh89227 A. thalia
40	146.5	11.6	164	8 ADG43962	Adg43962 A. thalia
41	144	11.4	169	6 AAE36077	Aae36077 Flax Conl
42	140.5	11.1	146	6 AAO27313	Aao27313 Brazil nu
43	140.5	11.1	146	8 ADN37294	Adn37294 Brazil nu
44	139	11.0	140	1 AAP91891	Aap91891 Brazil nu
45	139	11.0	155	2 AAW23587	Aaw23587 Mabinlin

#### ALIGNMENTS

RESULT 1

AAW24153  
ID AAW24153 standard; protein; 157 AA.

XX AC AAW24153;

XX DT 17-OCT-2003 (revised)

XX DT 29-DEC-1997 (first entry)

XX DE Peanut allergen Ara hII.

XX OS Peanut; seed storage protein; allergen; allergy; hypersensitivity;

KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;

KW ELISA; analysis; Ara hII.

XX OS Arachis hypogaea; strain Florunner.

XX WO97241139-A1.

XX PD 10-JUL-1997.

XX PF 23-SEP-1996; 96WO-US015222.

XX PR 29-DEC-1995; 95US-0009455P.

XX PR 04-MAR-1996; 96US-00610424.

XX PA (UYAR-) UNIV ARKANSAS.

XX PI Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;

XX DR WPI; 1997-363453/33.

XX DR N-PSDB; AAT76615.

XX PT Peanut allergens Ara hI and Ara hII - used for vaccination and in two-

```
PT site monoclonal antibody based ELISA.
XX
PS Claim 31; Page 198; 354pp; English.
XX
CC This polypeptide comprises major peanut allergen Ara hII. Its sequence
CC was deduced from a cDNA clone (AA776615) isolated from peanut seed cDNA
CC using a primer (see AAT76617) based on an isolated Ara hI peptide (see
CC AAW24151). The sequence shows significant homology with the conglutin
CC family of seed storage proteins of other legumes. The allergen is
CC recognised by serum IgE from a large proportion of individuals with
CC peanut hypersensitivity. Ara hII and Ara hI (see AAW24149-50) can be used
CC to raise monoclonal antibodies which are used in a specific two-site MAb
CC ELISA for the detection of Ara hI or Ara hII (claimed). IgE-binding Ara
CC hII antigen epitopes (see AAW24188-93) may be used in vaccines to protect
CC against allergic reactions to peanut allergens, e.g. anaphylactic shock.
CC (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 8-2e-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 2 Gaps: 0

US-10-728-051-2 (1-717) x AAW24153 (1-157)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCTGCTGCCACCGCATCTCGAGGCGAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACCTCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGGCGGCAACCTGAGG 121
DB 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGCGAGCAACATCTCATGCAGAGATCCCAACGTGACGAGGATTTCATATGAACGGGAC 181
DB 41 ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 241
DB 61 ProTyrSerProSerGlnAspProTyrSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAGG 301
DB 81 SerSerGlnHisGlnGluArgCysCysAsnGlnLeuAsnGluPheGluAsnGlnArg 100
QY 302 TGCATGTGCGAGGCGATTGCAACAGATCATGAGAACCCAGAGCGATAGGTTGCGGGGAGG 361
DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGAGCAAGTTCAAGAGGAGCTCAGGAACTTGCTCTCAACAGTCGCGCCTTAGG 421
DB 121 GlnGlnGluGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAAGGCTTGCGACTTGAGCTGAAAGTGGCGGCGAGACAGATAC 472
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 2
ID AAW24164
XX AAW24164 standard; protein; 157 AA.
AC AAW24164;
XX
XX 17-OCT-2003 (revised)
DT 29-DEC-1997 (first entry)
XX
DE Peanut allergen Ara hII.
XX Peanut; seed storage protein; allergen; allergy; hypersensitivity;
.KW
```

```
KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
KW ELISA; analysis; Ara hII.
XX
OS Arachis hypogaea; strain Florunner.
XX
PN WO9724139-A1.
XX
PD 10-JUL-1997.
XX
PF 23-SEP-1996; 96WO-US015222.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 04-MAR-1996; 96US-00610424.
XX
XX (UYAR-) UNIV ARKANSAS.
XX
PI Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;
XX
DR WPI; 1997-363453/33.
DR N-PSDB; AAT76614.
XX
PT Peanut allergens Ara hI and Ara hII - used for vaccination and in two-
PT site monoclonal antibody based ELISA.
XX
PS Claim 31; Page 219; 354pp; English.
XX
CC This polypeptide comprises major peanut allergen Ara hII. Its sequence
CC was deduced from cDNA clone P38 (AAT76614), isolated from peanut seed
CC cDNA using a primer (see AAT76617) based on an isolated Ara hI peptide
CC (see AAW24151). The sequence shows significant homology with the
CC conglutin family of seed storage proteins of other legumes. The allergen
CC is recognised by serum IgE from a large proportion of individuals with
CC peanut hypersensitivity. Ara hII and Ara hI (see AAW24149-50) can be used
CC to raise monoclonal antibodies which are used in a specific two-site MAb
CC ELISA for the detection of Ara hI or Ara hII (claimed). IgE-binding Ara
CC hII antigen epitopes (see AAW24188-93) may be used in vaccines to protect
CC against allergic reactions to peanut allergens, e.g. anaphylactic shock.
CC (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 8-2e-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 2 Gaps: 0

US-10-728-051-2 (1-717) x AAW24164 (1-157)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCTGCTGCCACCGCATCTCGAGGCGAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACCTCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGGCGGCAACCTGAGG 121
DB 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGCGAGCAACATCTCATGCAGAGATCCCAACGTGACGAGGATTTCATATGAACGGGAC 181
DB 41 ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 241
DB 61 ProTyrSerProSerGlnAspProTyrSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAGG 301
DB 81 SerSerGlnHisGlnGluArgCysCysAsnGlnLeuAsnGluPheGluAsnGlnArg 100
QY 302 TGCATGTGCGAGGCGATTGCAACAGATCATGAGAACCCAGAGCGATAGGTTGCGGGGAGG 361
```

Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120  
 QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAAGTTCCTCAACAGATCGCGCTTAGG 421  
 Db 121 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140  
 QY 422 GCACCAAGCGTTCGAGCTTCGAGCTCGAAGTGGCGGAGAGACAGATAC 472  
 Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157  
 RESULT 3  
 AAY15245  
 ID AAY15245 standard; protein; 157 AA.  
 XX AC AAY15245;  
 XX DT 17-OCT-2003 (revised)  
 XX DT 09-NOV-1999 (first entry)  
 DE Peanut allergen, Ara h 2, amino acid sequence.  
 XX KW allergy; immune response; transgenic; allergen; epitope;  
 KW immunoglobulin E; Ig E; binding site; peanut.  
 XX OS Arachis hypogaea.  
 XX PN WO9938978-A1.  
 XX PD 05-AUG-1999.  
 XX PF 29-JAN-1999; 99WO-US002031.  
 XX PR 31-JAN-1998; 98US-0073283P.  
 PR 13-FEB-1998; 98US-0074590P.  
 PR 13-FEB-1998; 98US-0074624P.  
 PR 13-FEB-1998; 98US-0074633P.  
 PR 27-AUG-1998; 98US-00141220.  
 XX (UYAR-) UNIV ARKANSAS.  
 PA (UYNY ) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.  
 PA (SOSI/) SOSIN H.  
 XX PI Sosin H, Bannon GA, Burks AW, Sampson HA;  
 XX WPI; 1999-479189/40.  
 DR N-PSDB; AA206383.  
 XX PT Modified allergen with reduced IgE binding, useful for treating e.g.  
 PT allergies.  
 XX PS Disclosure; Page 38; 46pp; English.  
 CC This is the amino acid sequence of the Ara h 2 protein from Arachis  
 CC hypogaea. The Ara h 2 protein has 10 IgE (immunoglobulin E) binding  
 CC epitopes, three of which are immunodominant (AAY15272, AAY15275, and  
 CC AAY15276). By modifying the IgE binding sites the ability of the allergen  
 CC to provoke an immune response is downregulated. The epitopes of the IgE  
 CC binding sites can therefore be modified in genetically engineered plants  
 CC and animals to elicit less of an allergic response. (Updated on 17-OCT-  
 CC 2003 to standardise OS field)  
 XX SQ Sequence 157 AA;  
 Alignment Scores:  
 Pred. No.: 8-2e-90 Length: 157  
 Score: 842.00 Matches: 157  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 66.56% Indels: 0  
 DB: 2 Gaps: 0  
 US-10-728-051-2 (1-717) x AAY15245 (1-157)

QY 2 CTCACACTACTAGTACCCCTCGCCCTTTTCCTCTCGCTGCCCGCATCTGCGAGGACG 61  
 Db 1 LeuThrIleLeuValAlaLeuAlaPheLeuLeuAlaHisAlaSerAlaArgGln 20  
 QY 62 CAGTGGAAATCCCAAGGAGACAGAGATGCGAGAGCCAGCTCCAGAGGGGGAACCTGAGG 121  
 Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40  
 QY 122 CCTCGGAGCAACATCTCATGCAAGAGATCCCAACGTGACGAGGATTCATATGAACGGAC 181  
 Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60  
 QY 182 CGGTACAGCCCTAGTACGATCCGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTGGA 241  
 Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80  
 QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTCAACGAGTTTCAGAACCAACAAAGG 301  
 Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100  
 QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGGAAGAACCCAGAGCGATAGTTGCGAGGAGG 361  
 Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120  
 QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAAGTTCCTCAACAGATCGCGCTTAGG 421  
 Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140  
 QY 422 GCACCAAGCGTTCGAGCTTCGAGCTCGAAGTGGCGGAGAGACAGATAC 472  
 Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157  
 RESULT 4  
 AAU04707  
 ID AAU04707 standard; protein; 157 AA.  
 XX AC AAU04707;  
 XX DT 23-OCT-2001 (first entry)  
 XX DE Anaphylactic antigen Ara h 2.  
 XX KW Ara h 2; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;  
 KW allergy; mast cell; basophil; mouse.  
 XX OS Mus sp.  
 XX PN WO200140264-A2.  
 XX PD 07-JUN-2001.  
 XX PF 06-DEC-2000; 2000WO-US033124.  
 XX PR 06-DEC-1999; 99US-00455294.  
 PR 23-JUN-2000; 2000US-0213765P.  
 PR 27-SEP-2000; 2000US-0235797P.  
 XX (PANA-) PANACEA PHARM LLC.  
 PA (UYAR-) UNIV ARKANSAS.  
 PA (MOUN ) MOUNT SINAI SCHOOL MEDICINE NEW YORK.  
 XX Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;  
 XX WPI; 2001-381378/40.  
 XX PT Antigenic fragments useful for reducing anaphylactic risk and reducing  
 PT the severity and/or number of allergic symptoms in individuals sensitive  
 PT to antigens, have reduced ability to bind Immunoglobulin E.  
 XX Claim 7; Fig 10; 100pp; English.  
 XX The sequence represents the amino acid sequence of anaphylactic antigen  
 CC Ara h 2. Ara h 2 is an anaphylactic antigen (A), which was used to design

CC antigenic peptides having a reduced ability to bind immunoglobulin E  
CC (IgE) as compared with the intact (A), or having a sequence substantially  
CC identical to a portion of sequence of an antigen that includes at least  
CC one IGE binding site, where at least one IGE binding site of the peptide  
CC is altered. The antigenic peptides are used in a composition which is  
CC useful for reducing risk or severity of allergic reaction to an antigen.  
CC This is done by identifying an individual at risk of allergic reaction to  
CC an antigen by identifying prior display of allergic symptoms when exposed  
CC to the antigen, or a familial relationship with an individual who  
CC previously displayed allergic symptoms when exposed to the antigen.  
CC Following this an antigen-specific IGE present on one or more mast cells  
CC or basophils in the individual's serum is identified. The individual is  
CC then contacted with a peptide corresponding to a portion of the antigen,  
CC which is selected, formulated, and delivered so that binding of the  
CC peptide to antigen-specific IGE is reduced as compared with IGE binding  
CC of intact antigen. The composition is also useful for treating and  
CC preventing allergic reactions

XX Sequence 157 AA;

Alignment Scores:  
Pred. No.: 8.2e-90 Length: 157  
Score: 842.00 Matches: 157  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 66.56% Indels: 0  
DB: 4 Gaps: 0

US-10-728-051-2 (1-717) x AAU04707 (1-157)

QY 2 CTACACATAGTACGCTCGCCCTTTCTCTCTGCTCCACGATCTGCGAGCGAG 61  
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20  
QY 62 CAGTGGGAATCCAAAGGAGACAGATGCCAGGCGAGCTCCAGAGGCGGACCTGAGG 121  
DB 21 GlnTrpGluLeuGlnGlyAspArgGlySerGlnLeuGluArgAlaAsnLeuArg 40  
QY 122 CCTGCGAGCAACATCTCATGCGAGAGATCCAACTGACGAGGATTCATATGAACGGGAC 181  
DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60  
QY 182 CGTACAGCCCTAGTCCGATCGGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241  
DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80  
QY 242 TCCTCTCAGCAACCAAGAGAGGTCTGCAATGAGCTGAACGAGTTTCAGAACCAACAGG 301  
DB 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100  
QY 302 TGCATGTGCGAGGCATTTGCAACAGATCATGGAAGCAGAGCGATAGGTTGCGAGGAGG 361  
DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120  
QY 362 CAACAGGACCAAGTTCAGAGGAGCTCAGGAATCTGCCTCAACAGTGCAGCGCTTAGG 421  
DB 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140  
QY 422 GCACCAAGGTTGCGACTTGGACGTCGAAAGTGGCGGAGACAGATAC 472  
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 5  
ABU52464

ID ABU52464 standard; protein; 157 AA.

XX AC ABU52464;

XX 10-MAR-2003 (first entry)

XX Peanut Ara h2 protein.

XX Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;

anaphylactic food allergen; antiallergenic; vaccine; wound healing.  
Arachis hypogaea.  
WO200274250-A2.  
26-SEP-2002.  
18-MAR-2002; 2002WO-US009108.  
16-MAR-2001; 2001US-0276822P.  
18-MAR-2002; 2002US-00276822.  
(PANA-) PANACEA PHARM.  
Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;  
Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;  
Rabjohn PA, Shin DS, Stanley JS;  
WPI; 2003-018765/01.  
N-PSDB; ABX70506.  
New modified anaphylactic food allergen, useful for preventing or  
treating allergic reactions associated with e.g. anaphylactic allergens.  
Claim 27; Fig 41; 300pp; English.

The invention relates to a modified anaphylactic food allergen has an  
amino acid sequence that is substantially identical to that of natural  
anaphylactic food allergen, except for a cysteine residue that has been  
modified so that it cannot participate in the disulphide bond. The  
modification may also comprise mutation of the IGE binding sites to  
reduce allergenicity. Also included are: (1) a method of making a  
modified anaphylactic food allergen; (2) a nucleotide molecule encoding  
or for causing a site specific mutation in the modified anaphylactic food  
allergen; (3) a transgenic plant or animal expressing the modified  
anaphylactic food allergen; (4) a method of treating an individual by  
reducing the clinical response to a natural anaphylactic food allergen;  
and an isolated fragment of peanut allergen Ara h 1. The modified  
anaphylactic food allergen is useful for preventing or treating allergic  
reactions associated with any natural allergen such as food, insect,  
rubber or preferably anaphylactic allergens. It is also useful for  
treating wounds in mammals such as bovine, canine, feline, caprine,  
ovine, porcine, murine or equine species. The present sequence is a  
peanut allergen (e.g. Ara h1, h2 or h3)

SQ Sequence 157 AA;

Alignment Scores:

Pred. No.: 8.2e-90 Length: 157  
Score: 842.00 Matches: 157  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 66.56% Indels: 0  
DB: 4 Gaps: 0

US-10-728-051-2 (1-717) x ABU52464 (1-157)

QY 2 CTCACATAGTACGCTCGCCCTTTCTCTCTGCTCCACGATCTGCGAGGCGAG 61  
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20  
QY 62 CAGTGGGAATCCAAAGGAGACAGATGTCAGAGCCAGCTCGAGAGGCGGACCTGAGG 121  
DB 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40  
QY 122 CCTGCGAGCAACATCTCATGCGAGAGATCCAACTGACGAGGATTCATATGAACGGGAC 181  
DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60  
QY 182 CGTACAGCCCTAGTCCGATCGGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241  
DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80

```
QY 242 TCCTCTCAGACCAAGAGAGGTTGCAATGAGCTGAAACGAGTTTGAGAACAAACGAGG 301
DB 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGGAGAACAGAGCGATAGTTGCAGGGGAGG 361
DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGCAGCCTTAGG 421
DB 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACACAGCGTTGCGACTTGGACGTGCGAAAGTGGCGGCGAGACAGATAC 472
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157

RESULT 6
ADG27518
ID ADG27518 standard; protein; 157 AA.
XX AC ADG27518;
DT 26-FEB-2004 (first entry)
XX DE Peanut allergen Ara h2.
XX KW Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycinin A2B1a; Jug n1;
KW antiallergic; vullnerary; anaphylactic food allergen; IGE; allergy; wound.
XX OS Arachis hypogaea.
XX PN US2003202980-A1.
XX PD 30-OCT-2003.
XX PF 18-MAR-2002; 2002US-00100303.
XX PR 29-DEC-1995; 95US-0009455P.
PR 23-SRP-1996; 96US-00717933.
PR 31-JAN-1998; 98US-0073283P.
PR 13-FEB-1998; 98US-0074590P.
PR 13-FEB-1998; 98US-0074624P.
PR 13-FEB-1998; 98US-0074633P.
PR 29-JUN-1998; 98US-00106872.
PR 27-AUG-1998; 98US-00141220.
PR 13-NOV-1998; 98US-00191593.
PR 29-JAN-1999; 99US-00240557.
PR 29-JAN-1999; 99US-00241101.
PR 11-FEB-1999; 99US-00248673.
PR 11-FEB-1999; 99US-00248674.
PR 02-MAR-1999; 99US-0122450P.
PR 02-MAR-1999; 99US-0122452P.
PR 02-MAR-1999; 99US-0122565P.
PR 02-MAR-1999; 99US-0122566P.
PR 11-MAR-1999; 99US-00267719.
PR 28-JAN-2000; 2000US-00494096.
PR 16-MAR-2001; 2001US-0276822P.
XX (CAPL/) CAPLAN M J.
PA (SOSI/) SOSIN H B.
PA (SAMP/) SAMPSON H.
PA (BANN/) BANNON G A.
PA (BURK/) BURKS A W.
PA (COCK/) COCKRELL G.
PA (COMP/) COMPADRE C M.
PA (CONN/) CONNAUGHTON C.
PA (HELM/) HELM R M.
PA (KING/) KING N E.
PA (KOPP/) KOPPER R A.
PA (MALE/) MALEKI S J.
PA (RABJ/) RABJOHN P A.
```

```
PA (SHIN/) SHIN D S.
PA (STAN/) STANLEY J S.
PI Caplan MJ, Sosin HB, Sampson H, Bannon GA, Burks AW, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX WPI; 2003-875632/81.
DR N-PSDB; ADG27517.
XX New modified anaphylactic food allergen comprising a cysteine residue
PT which has been modified so that it cannot participate in the disulfide
PT bond, useful for treating allergic reactions or wounds.
XX Claim 27; SEQ ID NO 63; 194pp; English.
XX The invention relates to a modified anaphylactic food allergen whose
CC amino acid sequence is substantially identical to that of a natural
CC anaphylactic food allergen. The natural anaphylactic food allergen
CC includes at least one cysteine residue that participates in a disulphide
CC bond when the natural anaphylactic food allergen is in its native
CC conformation, except that the cysteine residue has been modified so that
CC it cannot participate in the disulphide bond. Also included are a method
CC of making a modified anaphylactic food allergen, a nucleotide molecule
CC encoding a modified anaphylactic food allergen defined above, a
CC nucleotide molecule for causing a site specific mutation in a gene
CC encoding a natural anaphylactic food allergen, a transgenic plant or
CC animal expressing a modified anaphylactic food allergen defined above, a
CC method of treating an individual by reducing the clinical response to a
CC natural anaphylactic food allergen by administering a modified
CC anaphylactic food allergen and an isolated fragment of peanut allergen
CC Ara h 1, comprising at least 10 consecutive amino acids of ADG27464 or
CC ADG27465. About 10-17% of the amino acids have been modified in at least
CC one IGE epitope or all the IGE epitopes recognised when the natural
CC anaphylactic food allergen is contacted with serum IGE from individual(s)
CC allergic to the natural anaphylactic food allergen. The invention
CC discloses Peanut allergens Ara h1, Ara h2, Ara h3 (and their encoding
CC cDNAs), Soybean Glycinin A2B1a and IGE-binding epitopes of the English
CC walnut allergen Jug n1. The modified anaphylactic food allergen can be
CC used for treating allergic reactions or wounds. The present sequence
CC represents a Peanut allergen of the invention (or its fragment).
XX Sequence 157 AA;
SQ Alignment Scores:
Pred. No.: 8.2e-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 7
US-10-728-051-2 (1-717) x ADG27518 (1-157)
QY 2 CTCACCATACTAGTACCCCTCGCCCTTTCCTCTCTCGCTGCCACGCATCTCGAGCGAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAAGTCCCAAGGAGACAGAAAGTGCAGAGCCAGCTCGAGAGGGGGAACCTGAGG 121
DB 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCCTGCGAGCAACATCTCATGCAAGAGATCCAACTGACAGAGATTTCATATGAACCGGAC 181
DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCGGAGAGCGCTGGA 241
DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCCAATGAGCTGAACGAGTTTGTAGAACCAACCAAGG 301
DB 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
```

QY 302 TGCATGTGCGAGGATTGCAACAGATCATGGAGAACGAGCCGATAGTTGCGAGGAGG 361  
DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120  
QY 362 CAACAGGAGCAACAGTTCAGAGGGAGCTCAGGAATTCCTCAACAGTCGGGCTTAGG 421  
DB 121 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140  
QY 422 GCACCAAGGCTTGGGAGTTCGAAAGTGGCGGCGAGACAGATAC 472  
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157  
RESULT 7  
ID AAB82383 standard; protein; 207 AA.  
XX  
AC AAB82383;  
XX  
DT 11-SEP-2003 (revised)  
XX 23-JUL-2001 (first entry)  
XX  
DE Peanut allergen Ara h2 gene product.  
XX  
KW Peanut; allergen; Ara h2; transgenic plant; allergy.  
XX  
OS Arachis hypogaea.  
XX  
FH Location/Qualifiers  
FT Peptide 1..21  
FT Protein /label= Signal\_peptide  
FT /label= Mature\_protein  
XX  
PN WO200136621-A2.  
XX  
XX  
XX 25-MAY-2001.  
XX  
XX 20-NOV-2000; 2000WO-US031657.  
XX  
XX 19-NOV-1999; 99US-0167255P.  
XX  
XX (UVAL-) UNIV ALABAMA A & M.  
XX  
XX Dodo HW, Arntzen CJ, Konan KN, Viquez OM;  
XX  
XX WPI; 2001-355630/37.  
XX  
XX N-PSDB; AAF90336, AAF90337.  
XX  
XX Producing transgenic peanut plants that produce allergen-free seeds,  
XX useful in non-allergenic foods, by antisense or sense co-suppression of  
XX allergen-encoding genes.  
XX  
XX Example 1; Fig 2; 72pp; English.  
XX  
XX The present sequence is that of the peanut allergenic protein (AP)  
XX encoded by the Ara h2 gene (see AAF90336). The invention relates to a  
XX method for producing a peanut plant having reduced, or undetectable, AP  
XX content in its seed. A peanut plant cell is transformed with a DNA  
XX construct containing an antisense AP gene and/or sense AP gene, or their  
XX fragments, regenerated to plants, and fertile transgenic plants that  
XX produce seeds with reduced AP content are identified. The AP sense or  
XX antisense gene may comprise at least a portion of the Ara h2 gene  
XX sequence. The seeds are useful for preparation of allergen-free foods.  
XX Recombinant AP may be produced and used to produce antibodies useful for  
XX detecting AP in foods, and for treatment or prevention of peanut allergy.  
XX (Updated on 11-SEP-2003 to standardise OS field)

Alignment Scores: 2.39e-89 Length: 207  
Pred. No.: 838.50 Matches: 162  
Score:

Percent Similarity: 90.27% Conservative: 5  
Best Local Similarity: 87.57% Mismatches: 17  
Query Match: 66.28% Indels: 1  
DB: 4 Gaps: 1  
US-10-728-051-2 (1-717) x AAB82383 (1-207)  
QY 2 CTCACCATAGTAGCCCTCGCCCTTTTCCTCCTCGTCCCAACGATCTGCGAGGCAG 61  
DB 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHlaHlaSerAlaArgGln 23  
QY 62 CAGTGGGAATCTCAAGGAGACAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAGG 121  
DB 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43  
QY 122 CCTCGGAGCAACATCTCATGCGAGAGATCCAACTGAGAGAGATTTCATATGACGGGAC 181  
DB 44 ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp 63  
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATATGCGAGAGCGCTGGA 241  
DB 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 83  
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGACGAGTTTTCAGAACCAAGAG 301  
DB 84 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 103  
QY 302 TGCATGTGCGAGGATTGCAACAGATCATGGAGAACGAGCCGATAGTTGCGAGGAGG 361  
DB 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123  
QY 362 CAACAGGAGCAACAGTTCAGAGGGAGCTCAGGAATTCCTCAACAGTCGGGCTTAGG 421  
DB 124 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 143  
QY 422 GCACCAAGGCTTGGGAGTTCGAAAGTGGCGGCGAGACAGATAC 481  
DB 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgProArgIleProPro 163  
QY 482 ATCTCAAAAAAGAAAAAGAAAAAGAAAAAGTATATATAAGC---TATTATCTA 538  
DB 164 IleLeuThrGlySerArgSerArgHisGlnSerProTyrGlyAsnArgArgTyrSer 183  
QY 539 TGGTTATGTTAGTT 553  
DB 184 AlaMetCysLeuLeu 188

RESULT 8  
ID AAY40973 standard; protein; 156 AA.  
XX  
AC AAY40973;  
XX  
DT 17-OCT-2003 (revised)  
DT 06-DEC-1999 (first entry)  
XX  
DE Ara h 2 protein fragment.  
XX  
KW Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;  
KW allergic reaction; Ara h 2.  
XX  
OS Arachis hypogaea.  
XX  
PN WO9945961-A1.  
XX  
XX 16-SEP-1999.  
XX  
XX 12-MAR-1999; 99WO-US005494.  
XX  
XX 12-MAR-1998; 98US-0077763P.  
XX 11-MAR-1999; 99US-0007763.  
XX (UYAR-) UNIV ARKANSAS.



XX  
PI Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;  
PI Sampson H, Compadre CM, Huang SK, Maleki SU, Kopper RA;  
XX WPI; 1999-551218/46.  
XX  
PT Tertiary structure of peanut allergen Ara h 1 for protection of a host  
PT animal from allergic reaction.  
XX  
PS Disclosure; Page 104; 193pp; English.  
XX  
CC The invention provides a tertiary structure for the peanut allergen Ara H  
CC 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding  
CC epitopes. The invention also provides an isolated recombinant peanut  
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut  
CC allergen Ara h 3. Molecules of the invention are used to protect a host  
CC animal from allergic reaction, particularly using a modified allergen  
CC which is less reactive with IgE. The invention may also be used to ensure  
CC that the allergen is not introduced into genetically modified food. The  
CC present sequence represents an Ara h 2 protein fragment. (Updated on 17-  
CC OCT-2003 to standardise OS field)  
XX  
SQ Sequence 156 AA;  
  
Alignment Scores:  
Pred. No.: 2.13e-88 Length: 156  
Score: 830.00 Matches: 155  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 65.61% Indels: 0  
DB: 2 Gaps: 0  
  
US-10-728-051-2 (1-717) x AAY40973 (1-156)  
QY 2 CTCACTACTAGTAGGCTCGCCCTTTCTCTCTCTGCCACGCGATCGGAGGCGAG 61  
Db 1 LeuThrlleuValalaLeuAlaLeuPheleuLeuAlaHisalaSerAlaArgGln 20  
  
QY 62 CAGTGGGAATCCACAGAGACAGAAGATGCCAGAGCGAGCTCAGAGGGCGAACCCTGAGG 121  
Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40  
  
QY 122 CCTCGGAGCAACATCTCATCGCAGAGATCCACAGTGCACAGGATTCATATGACGGGAC 181  
Db 41 ProCysGluGlnHisLeuMetGlnLysileGlnArgAspGluAspSerTyrGluArgAsp 60  
  
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGGCTGGA 241  
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80  
  
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTGACAGCTGAACAGAGTTTGAGAACCAACGAGG 301  
Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100  
  
QY 302 TGCATGTCGAGGATTCGACAGATCATGAGACAGAGCGATAGTTGCGAGGGAGG 361  
Db 101 CysMetCysGluAlaLeuGlnGlnMetGluAsnGlnSerAspArgLeuGlnGlyArg 120  
  
QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTCCGCGCTTAGG 421  
Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140  
  
QY 422 GCACACAGGCTTGCGACTTGGACGTGCGAAGTGGCGGCGAGAC 466  
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAsp 155  
  
RESULT 9  
ADO38314  
ID ADO38314 standard; peptide; 156 AA.  
XX  
AC ADO38314;  
XX  
DT 15-JUL-2004 (first entry)

XX  
DE Ara h 2 MHC-class II-presented epitope #1.  
XX  
KW Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;  
KW Antirheumatic; Neuroprotective; Antiinflammatory; Dermatological;  
KW Immunosuppressive; Antidiabetic; Antichyroid; Antiasthmatic;  
KW Allergic; Cytostatic; Antipsoriatic; Gene Therapy; Vaccine;  
KW MHC Class II; Ii-Key motif; immune response; anthrax; EBOLA; HIV;  
KW influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;  
KW rickettsia; rheumatoid arthritis; diabetes mellitus; myasthenia gravis;  
KW lupus erythematosus; diabetes mellitus; dermatomyositis; pemphigus; asthma;  
KW autoimmune thyroiditis; scleroderma; colitis; cancer; psoriasis;  
KW allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;  
KW adenoma; peanut; Ara h 2.  
XX  
OS Arachis hypogaea.  
XX  
XX US2004058881-A1.  
XX  
PN 25-MAR-2004.  
PD  
XX 24-SEP-2002; 2002US-00253286.  
PF  
XX 24-SEP-2002; 2002US-00253286.  
PR  
XX (ANTI-) ANTIGEN EXPRESS INC.  
PA  
XX Humphreys RE, Xu M;  
PI  
XX WPI; 2004-294259/27.  
DR  
XX  
XX New non-naturally occurring protein or polypeptide modified by  
PT recombinant DNA techniques, useful for treating multiple sclerosis,  
PT diabetes mellitus, myasthenia gravis, scleroderma, allergic rhinitis,  
PT colitis, cancer or psoriasis.  
XX  
PS Example 2; Page 18; 90pp; English.  
XX  
CC The invention relates to a non-naturally occurring protein or polypeptide  
CC (I) modified by recombinant DNA techniques comprising: a C-terminal  
CC element comprising an MHC Class II-presented epitope; an N-terminal  
CC element comprising an Ii-key motif; and an intervening element comprising  
CC a sequence of 4-11 amino acid residues where the modification by  
CC recombinant DNA techniques taking place within elements (b) and (c). Also  
CC described are methods for: suppressing or enhancing an immune response  
CC directed toward an MHC (major histocompatibility complex) Class II-  
CC presented epitope of interest. Suppressing an immune response directed  
CC toward an MHC Class II-presented epitope of interest comprises: providing  
CC a nucleic acid sequence encoding the MHC Class II-presented epitope of  
CC interest, the nucleic acid sequence encoding an Ii-key motif located 4-11  
CC amino acids upstream from the N-terminal residue of the MHC Class II-  
CC presented epitope of interest; and modifying the Ii-key motif to decrease  
CC its conformance to the archetypal Ii-Key regulatory motif. Enhancing an  
CC immune response directed toward an MHC Class II-presented epitope of  
CC interest comprises: providing a nucleic acid sequence encoding the MHC  
CC Class II-presented epitope of interest, the nucleic acid sequence lacking  
CC an Ii-key motif located 4-11 amino acids upstream from the N-terminal  
CC residue of the MHC Class II-presented epitope of interest; and modifying  
CC the nucleic acid sequence to introduce an Ii-key motif appropriately  
CC spaced from the MHC Class II-presented epitope. The protein or  
CC polypeptide of interest corresponds to a protein or polypeptide encoded  
CC by an infectious pathogen selected from anthrax, EBOLA, HIV or influenza,  
CC preferably vaccinia virus. The non-naturally occurring protein or  
CC polypeptide (I) modified by recombinant DNA techniques is useful for  
CC treating infectious diseases caused or associated with infection by a  
CC bacterium, virus, parasite, fungus, rickettsia or other infectious  
CC agents. It is also useful for treating rheumatoid arthritis, multiple  
CC sclerosis, lupus erythematosus, diabetes mellitus, myasthenia gravis,  
CC autoimmune thyroiditis, scleroderma, dermatomyositis, pemphigus, asthma,  
CC allergic rhinitis, topical dermatitis, colitis, cancer, psoriasis or  
CC adenomas. The present sequence represents the amino acid sequence of a  
CC peanut allergen Ara h 2 MHC class II-presented epitope used in the  
CC invention.

```
XX SQ Sequence 156 AA;
Alignment Scores:
Pred. No.: 4.23e-87 Length: 156
Score: 819.00 Matches: 153
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.74% Indels: 0
DB: 8 Gaps: 0

US-10-728-051-2 (1-717) x ADO38314 (1-156)
Qy 2 CTCACCATAGTAGAGCCCTCGCCCTTTCTCTCGCTGCCACGGCATCTGGAGGCAG 61
Db 4 LeuThrileLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
Qy 62 CAGTGGGAATCCNAGGAGACAGAGATGTCAGAGCGCTCAGAGGGCGAACTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
Qy 122 CCTGCGGAGCAATCTCATGCGAGAGATCCACGTCACGAGGATTCATATGACGGGAC 181
Db 44 ProCysGlnGlnHisLeuMetGlnLysileGlnArgAspGluAspSerTyrGluArgAsp 63
Qy 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATATCATCGAGGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
Qy 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAGCTGAGCTTTGAGAACACCAAAG 301
Db 84 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
Qy 302 TGCATGTGCGAGGATTCGACAGATCATGAGACCCAGATCATGATCGAGGCGCTGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGlnAsnGlnSerAspArgLeuGlnGlyArg 123
Qy 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGGCTTAGG 421
Db 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
Qy 422 GCACCAAGCTGGAGCTGGAGCTGGAAGTGGCGGC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156

RESULT 10
ID ABU52576
AC ABU52576 standard; protein; 157 AA.
XX ABU52576;
DT 10-MAR-2003 (first entry)
XX Peanut Ara h2 mutant 5 (Q20A/Q31A/D60A/D67A).
DE Allergy; Ara h1; Ara h2; Ara h3; IGE binding site; peanut; mutant;
KW mutein; anaphylactic food allergen; anti-allergenic; vaccine;
KW wound healing.
XX Homo sapiens.
OS
XX WO200274250-A2.
FN
XX 26-SEP-2002.
PD
XX 18-MAR-2002; 2002WO-US009108.
PF
XX 16-MAR-2001; 2001US-0276822P.
PR
XX 18-MAR-2002; 2002US-00276822.
XX (PANA-) PANACEA PHARM.
PA
XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI
```

PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;  
PI Rabjohn PA, Shin DS, Stanley JS;  
XX WPI; 2003-018765/01.

XX New modified anaphylactic food allergen, useful for preventing or  
PT treating allergic reactions associated with e.g. anaphylactic allergens.  
XX Example 13; Page: 300pp; English.

XX The invention relates to a modified anaphylactic food allergen has an  
CC amino acid sequence that is substantially identical to that of natural  
CC anaphylactic food allergen, except for a cysteine residue that has been  
CC modified so that it cannot participate in the disulphide bond. The  
CC modification may also comprise mutation of the IGE binding sites to  
CC reduce allergenicity. Also included are: (1) a method of making a  
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding  
CC or for causing a site specific mutation in the modified anaphylactic food  
CC allergen; (3) a transgenic plant or animal expressing the modified  
CC anaphylactic food allergen; (4) a method of treating an individual by  
CC reducing the clinical response to a natural anaphylactic food allergen;  
CC and an isolated fragment of peanut allergen Ara h 1. The modified  
CC anaphylactic food allergen is useful for preventing or treating allergic  
CC reactions associated with any natural allergen such as food, insect,  
CC rubber or preferably anaphylactic allergens. It is also useful for  
CC treating wounds in mammals such as bovine, canine, feline, caprine,  
CC ovine, porcine, murine or equine species. The present sequence is a food  
CC allergen, mutated to alter its IGE binding characteristics. Note: The  
CC present sequence is not shown in the specification but was created by the  
CC indexer using information provided in the specification  
XX Sequence 157 AA;

Alignment Scores:  
Pred. No.: 1.65e-86 Length: 157  
Score: 814.00 Matches: 153  
Percent Similarity: 97.45% Conservative: 0  
Best Local Similarity: 97.45% Mismatches: 4  
Query Match: 64.35% Indels: 0  
DB: 6 Gaps: 0

US-10-728-051-2 (1-717) x ABU52576 (1-157)

Qy 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGGCATCTCGAGGCAG 61  
Db 1 LeuThrileLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgAla 20  
Qy 62 CAGTGGGAATCCAAAGGAGACAGAGATGCCAGACCCAGCTCGAGAGGCGCACTGAGG 121  
Db 21 GlnTrpGluLeuGlnGlyAspArgCysAlaSerGlnLeuGluArgAlaAsnLeuArg 40  
Qy 122 CCTGCGGAGCAATCTCATGCGAGAGATCCACGTCACGAGGATTCATATGACGGGAC 181  
Db 41 ProCysGlnGlnHisLeuMetGlnLysileGlnArgAspGluAspSerTyrGluArgAla 60  
Qy 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGCTGGA 241  
Db 61 ProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArgArgGlyAlaGly 80  
Qy 242 TCCTCTCAGCCACCAAGAGAGGTGTGCAATGAGCTGAAACGAGTTTGAGAACACCAAAGG 301  
Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100  
Qy 302 TGCATGTGCGAGGATTCGACAGATCATGAGACCCAGATCATGATCGAGGCGGAGG 361  
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGlnAsnGlnSerAspArgLeuGlnGlyArg 120  
Qy 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGGCTTAGG 421  
Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140  
Qy 422 GCACCAAGCTGGAGCTGGAGCTGGAAGTGGCGGCAGACAGATAC 472

141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

Db  
RESULT 11  
ADM12096  
ID ADM12096 standard; protein; 156 AA.  
XX  
AC ADM12096;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Arachis hypogaea 2 (Ara h2) protein.  
XX  
KW antigen presentation enhancing hybrid polypeptide; mammalian li-key;  
KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;  
KW antiarthritic; neuroprotective; dermatological; immunosuppressive;  
KW antiinflammatory; antidiabetic; antithyroid; immune;  
KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;  
KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;  
KW scleroderma; dermatomyositis; pemphigus.  
XX  
OS Arachis hypogaea.  
XX  
PN US2003235594-A1.  
XX  
PD 25-DEC-2003.  
XX  
PP 17-SEP-2002; 2002US-00245871.  
XX  
PR 14-SEP-1999; 99US-00396813.  
PR 17-JUL-2002; 2002US-00197000.  
XX  
PA (ANTI-) ANTIGEN EXPRESS INC.  
XX  
PI Humphreys R, Xu M;  
XX  
DR WPI; 2004-070554/07.  
XX  
PT Novel Ii-Key/antigen presentation enhancing hybrid polypeptide, useful  
PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus  
PT erythematosus and diabetes mellitus.  
XX  
XX Example 2; Page 19; 87pp; English.  
XX  
CC The invention relates to a novel antigen presentation enhancing hybrid  
CC polypeptide. The novel polypeptide has an N-terminal element consisting  
CC of 4-16 residues of a mammalian Ii-Key peptide and its non-N-terminal  
CC deletion modifications, a chemical structure covalently linking the N-  
CC terminal element to an MHC class II-presented epitope of a C-terminal  
CC element. The C-terminal element comprises an antigenic epitope, which  
CC binds to an antigenic peptide binding site of an MHC class II molecule.  
CC The antigen presentation enhancing hybrid polypeptide has the following  
CC activities: antibacterial, virucide, fungicide, antirheumatic,  
CC antiarthritic, neuroprotective, dermatological, immunosuppressive,  
CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation  
CC enhancing hybrid polypeptide is useful for modulating the immune response  
CC in an individual and for treating infections (such as bacteria, virus,  
CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus  
CC erythematosus, diabetes mellitus, myasthenia gravis, autoimmune  
CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence  
CC represents a mammalian Ii key related protein of the invention.  
XX  
SQ Sequence 156 AA;  
Alignment Scores:  
Pred. No.: 2.83e-86 Length: 156  
Score: 812.00 Matches: 152  
Percent Similarity: 99.35% Conservative: 0  
Best Local Similarity: 99.35% Mismatches: 1  
Query Match: 64.19% Indels: 0  
DB: 8 Gaps: 0  
US-10-728-051-2 (1-717) x ADM12096 (1-156)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTTCCTCCTCGCTCCCGACCGCATCTGCGAGCAG 61  
DB |||||  
4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23  
QY 62 CAGTGGAACTCCACGAGACAGAAAGATGCGAGCGAGCTCGAGAGGGGGAACCTCAGG 121  
DB |||||  
24 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43  
QY 122 CCTTGGAGCAACATCTCATGCAAGATCCAAAGTCCAGCGAGGATTTCATATGAACGGGAC 181  
DB |||||  
44 ProCysGluGlyHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63  
QY 182 CCTAGAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGCGAGAGCGCTGGA 241  
DB |||||  
64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83  
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACAGTTTGTAGAACCAACAAAGG 301  
DB |||||  
84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103  
QY 302 TGCATGTGCGAGGCGATTGCAACAGATCATGAGAACCGAGAGCGATAGTTGCGAGGGAGG 361  
DB |||||  
104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123  
QY 362 CAACAGAGCAACAGTTCAGAGCGGAGCTCAGGAAGTTCCTCAACAGTGGCGCTTAGG 421  
DB |||||  
124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143  
QY 422 GCACCAAGCGTTGCGACTTGGACGTCGAAAGTGGCGGC 460  
DB |||||  
144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156  
RESULT 12  
ABU52575  
ID ABU52575 standard; protein; 157 AA.  
XX  
AC ABU52575;  
XX  
DT 10-MAR-2003 (first entry)  
XX  
DE Peanut Ara h2 mutant 4 (E35A/P41A/D60A/D67A).  
XX  
KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;  
KW mutein; anaphylactic food allergen; antiallergenic; vaccine;  
KW wound healing.  
OS Homo sapiens.  
XX  
PN WO200274250-A2.  
XX  
PD 26-SEP-2002.  
XX  
PF 18-MAR-2002; 2002WO-US009108.  
XX  
PR 16-MAR-2001; 2001US-0276822P.  
PR 18-MAR-2002; 2002US-00276822.  
XX  
PA (PANA-) PANACEA PHARM.  
XX  
PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;  
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;  
PI Rabjohn PA, Shin DS, Stanley JS;  
XX  
DR WPI; 2003-018765/01.  
XX  
PT New modified anaphylactic food allergen, useful for preventing or  
PT treating allergic reactions associated with e.g. anaphylactic allergens.  
XX  
PS Example 13; Page; 300pp; English.  
XX  
CC The invention relates to a modified anaphylactic food allergen has an  
CC amino acid sequence that is substantially identical to that of natural  
CC anaphylactic food allergen, except for a cysteine residue that has been



||||| 61 ProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArgArgGlyAlaGly 80  
QY 242 TCCTCTCAGCACCAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAAACCAAGG 301  
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100  
QY 302 TGATGTGCGAGGATTCGCAACAGATCATGAGAACCCAGAGCGATAGGTTGCGAGGAGG 361  
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyAla 120  
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGGAGCTCAGGAACCTTGCTCAACAGTCGCGCTTAGG 421  
Db 121 GlnGlnGlnGlnPheLysArgGluAlaArgAsnLeuProGlnGlnCysGlyLeuArg 140  
QY 422 GCACCACAGGTTGCGACTTGGACGTCGAAAGTGGCGGCAGAGACAGATAC 472  
Db 141 AlaProGlnArgCysAspAlaAspValGluSerGlyGlyArgAspArgTyr 157  
RESULT 14  
ID AAY40968 standard; protein; 166 AA.  
XX AC AAY40968;  
XX 17-OCT-2003 (revised)  
DT 06-DEC-1999 (first entry)  
XX XX  
XX Recombinant Ara h 2 protein sequence.  
XX Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;  
KW allergic reaction; Ara h 2.  
XX Arachis hypogaea.  
XX WO9945961-A1.  
XX 16-SEP-1999.  
XX 12-MAR-1999; 99WO-US005494.  
XX 12-MAR-1998; 98US-0077763P.  
PR 11-MAR-1999; 99US-00077763.  
XX (UYAR-) UNIV ARKANSAS.  
XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;  
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;  
XX WPI; 1999-551218/46.  
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host  
PT animal from allergic reaction.  
XX Disclosure; Page 81; 193pp; English.  
XX The invention provides a tertiary structure for the peanut allergen Ara H  
CC 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding  
CC epitopes. The invention also provides an isolated recombinant peanut  
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut  
CC allergen Ara h 3. Molecules of the invention are used to protect a host  
CC animal from allergic reaction, particularly using a modified allergen  
CC which is less reactive with IgE. The invention may also be used to ensure  
CC that the allergen is not introduced into genetically modified food. The  
CC present sequence represents a recombinant Ara h 2 protein. (Updated on 17  
CC -OCT-2003 to standardise OS field)  
XX SQ Sequence 166 AA;  
Alignment Scores:  
Pred. No.: 2e-81 Length: 166  
Score: 771.00 Matches: 141  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 60.95% Indels: 0  
DB: 2 Gaps: 0  
US-10-728-051-2 (1-717) x AAY40968 (1-166)  
QY 50 TCTGCCAGCAGCAGTGGGNACTCCAAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAG 109  
Db 15 SerAlaArgGlnGlnTrpGluLeuGlnGlnIleArgCysGlnSerGlnLeuGluArg 34  
QY 110 GCCAACTCAGGCGCTTCGAGCAACATCTCATCGAGAAGATCCAACTGACGAGGATTCA 169  
Db 35 AlaAsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSer 54  
QY 170 TATGAACGGGACCCGTACAGCCCTAGTACGATCCGTACAGCCCTAGTCCATATGATCGG 229  
Db 55 TyrGluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg 74  
QY 230 AGAGGCGCTGGATCCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTGAG 289  
Db 75 ArgGlyAlaGlySerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGlu 94  
QY 290 AACAAACAAAGGTGTCATGTCGAGGCGATTGCAACAGATCATGAGAACCAAGAGCGATAGG 349  
Db 95 AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg 114  
QY 350 TTCAGGCGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAG 409  
Db 115 LeuGlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGln 134  
QY 410 TCGCGGCTTAGGCGCACACAGCGTTGCGACTTGGAGCTGCAAAAGTGGCGGCAGAGACAG 469  
Db 135 CysGlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg 154  
QY 470 TAC 472  
Db 155 Tyr 155  
RESULT 15  
ID ABUS2482 standard; protein; 166 AA.  
XX AC ABUS2482;  
XX 10-MAR-2003 (first entry)  
DT Peanut Ara h2 with N-terminal a T7 tag and a C-terminal His tag.  
DE Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;  
XX anaphylactic food allergen; anti-allergenic; vaccine; wound healing.  
KW Arachis hypogaea.  
OS Synthetic.  
OS Enterobacteria phage T7.  
XX WO200274250-A2.  
XX 26-SEP-2002.  
XX 18-MAR-2002; 2002WO-US009108.  
XX 16-MAR-2001; 2001US-0276822P.  
PR 18-MAR-2002; 2002US-00276822.  
XX (PANA-) PANACEA PHARM.  
XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;  
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;  
PI Rabbjohn PA, Shin DS, Stanley JS;  
XX WPI; 2003-018765/01.  
XX New modified anaphylactic food allergen, useful for preventing or

PT treating allergic reactions associated with e.g. anaphylactic allergens.

XX Example 13; Fig 51; 300pp; English.

XX The invention relates to a modified anaphylactic food allergen has an amino acid sequence that is substantially identical to that of natural anaphylactic food allergen, except for a cysteine residue that has been modified so that it cannot participate in the disulphide bond. The modification may also comprise mutation of the IGE binding sites to reduce allergenicity. Also included are: (1) a method of making a modified anaphylactic food allergen; (2) a nucleotide molecule encoding or for causing a site specific mutation in the modified anaphylactic food allergen; (3) a transgenic plant or animal expressing the modified anaphylactic food allergen; (4) a method of treating an individual by reducing the clinical response to a natural anaphylactic food allergen; and an isolated fragment of peanut allergen Ara h 1. The modified anaphylactic food allergen is useful for preventing or treating allergic reactions associated with any natural allergen such as food, insect, rubber or preferably anaphylactic allergens. It is also useful for treating wounds in mammals such as bovine, canine, feline, caprine, ovine, porcine, murine or equine species. The present sequence is a peanut allergen, Ara h2, with an N-terminal T7 tag and a C-terminal His tag

SQ Sequence 166 AA;

#### Alignment Scores:

Pred. No.:	2e-81	Length:	166
Score:	771.00	Matches:	141
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	60.95%	Indels:	0
DB:	6	Gaps:	0

US-10-728-051-2 (1-717) x ABUS2482 (1-166)

Qy	50	TCTCGAGGAGGAGTGGGAACTCCAGAGAGAGAGATGCCAGAGCCAGCTCGAGAGG	109
Db	15	SerAlaArgGlnGlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArg	34
Qy	110	GCGAACTGAGGGCCCTCGAGCAACATCTCATGAGAGATCCAACTGACGAGGATTCA	169
Db	35	AlaAsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSer	54
Qy	170	TATCAACGGGACCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGG	229
Db	55	TyrGluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg	74
Qy	230	AGAGCGCTGATCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGACGAGTTTGAG	289
Db	75	ArgGlyAlaGlySerSerGlnHisGlnGluArgCysAsnGlnLeuAsnGluPheGlu	94
Qy	290	AACAACCAAGGTGCATGTGCGAGGCGATTGCAACAGATCATGGAGAACAGAGCGATAGG	349
Db	95	AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg	114
Qy	350	TTGCGAGGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAG	409
Db	115	LeuGlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGln	134
Qy	410	TGCGGCTTAGGGCCACCACAGCGCTTGGACCTTGGACCTTGGACCTTGGACCTTGGAC	469
Db	135	CysGlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArg	154
Qy	470	TAC 472	
Db	155	Tyr 155	

Search completed: April 8, 2005, 05:01:20  
Job time : 102.896 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 8, 2005, 04:05:21 ; Search time 24.3307 Seconds  
(without alignments)  
5670.817 Million cell updates/sec

Title: US-10-728-051-2  
Perfect score: 1265  
Sequence: 1 gctcacatagtagtgcgc.....taaaagatcatgttttggtt 717

Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ n2p.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_P/US10728051/runat\_07042005\_125045\_14060/app\_query.fasta\_1.4757  
-DB=PIR -QPMT-fastan -SUFFIX=xpr -MINMATCH=0.1 -LOP=0.1 -LOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEA=500 -MINLEN=0 -MAXLEN=2000000000  
-USR=US10728051@cgn 1.1 228 @runat\_07042005\_125045\_14060 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	278	22.0	153	2 A33090	conglutin delta pr
2	243	19.2	158	1 T05710	2S albumin precurs
3	187.5	14.8	80	2 B23617	conglutin delta-2
4	184.5	14.6	323	2 S38887	2S albumin - commo
5	173.5	13.7	141	2 T10257	2S albumin precurs
6	154	12.2	154	2 S14947	2S albumin - Brazi
7	151	11.9	295	2 S01062	2S seed storage pr
8	147	11.6	186	2 A29802	napin precursor (g
9	146.5	11.6	164	1 NWMU1	2S albumin 1 precu
10	140.5	11.1	146	2 S14946	2S seed storage pr
11	139	11.0	155	2 JC5379	mablin II precur
12	139	11.0	258	1 RZCS	2S seed storage pr
13	132.5	10.5	165	2 T09252	seed storage prote
14	129	10.2	170	2 T08012	2S seed storage pr

15	128	10.1	172	2 S18871	2S-like storage pr
16	126.5	10.0	70	2 A59346	seed storage prote
17	125	9.9	162	2 T08013	2S seed storage pr
18	124.5	9.8	162	2 S49259	albumin 4 - easter
19	122.5	9.7	162	2 T08010	2S seed storage pr
20	121	9.6	100	2 S48180	mablin IV - Yunn
21	121	9.6	174	2 P80425	napin B33 precursor
22	119.5	9.4	180	2 S52025	napin (clones BNMN
23	118.5	9.4	106	2 S26636	napin n1b - rape
24	118.5	9.4	110	2 S20350	napin n1a - rape
25	118	9.3	104	2 S48176	mablin I-1 - Yun
26	117	9.2	104	2 S48178	mablin III - Yun
27	115	9.1	152	2 P80427	napin AHI precursor
28	111.5	8.8	173	2 T08011	2S seed storage pr
29	108.5	8.6	178	2 S07828	napin B - rape
30	108.5	8.6	178	2 S25127	2S storage protein
31	107	8.5	139	2 T09850	albumin 2S storage
32	107	8.5	139	2 T09878	albumin 2S storage
33	106.5	8.4	178	1 NWRP2	napin 2 precursor
34	106.5	8.4	178	2 A25997	napin precursor (n
35	105.5	8.3	164	1 NWMU3	2S albumin 3 precu
36	105.5	8.3	180	2 S10018	napin (clone BgNA
37	105.5	8.3	1099	2 A56155	tumor suppressor p
38	105.5	8.3	1909	2 A45592	liver stage antige
39	104.5	8.3	1263	2 T15496	hypothetical prote
40	104	8.2	145	2 S65479	allergen Sin a I (
41	103.5	8.2	166	1 NWMU4	2S albumin 4 precu
42	103.5	8.2	178	2 S25130	2S storage protein
43	102.5	8.1	178	2 S25134	2S storage protein
44	101.5	8.0	178	2 S25137	2S storage protein
45	101	8.0	145	2 PCL246	Sin a I allergen 1

ALIGNMENTS

RESULT 1

A33090  
conglutin delta precursor - narrow-leaved blue lupine  
C/Species: Lupinus angustifolius (narrow-leaved blue lupine)  
C/Date: 03-Aug-1990 #sequence\_revision 03-Aug-1990 #text\_change 09-Jul-2004  
C/Accession: S12404; A33090  
R/Gayler, K.R.; Kolivas, S.; Macfarlane, A.J.; Lilley, G.G.; Baldi, M.; Blagrove, R.J.;  
Plant Mol. Biol. 15, 879-893, 1990  
A/Title: Biosynthesis, cDNA and amino acid sequences of a precursor of conglutin delta,  
A/Reference number: S12404; MUID:91355912; PMID:2103479  
A/Accession: S12404  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-153 <GAY>  
A/Cross-references: UNIPROT:Q99235; GB:X53523; NID:g19140; PIDN:CAA37598.1; PID:g19141  
C/Superfamily: soybean 2S albumin  
C/Keywords: seed  
F/1-22/Domain: signal sequence #status predicted <SIG>  
F/23-55/Product: conglutin small chain #status predicted <MAT1>  
F/73-153/Product: conglutin large chain #status predicted <MAT2>

Alignment Scores:

Pred. No.: 3 84e-18 Length: 153  
Score: 278.00 Matches: 62  
Percent Similarity: 55.00% Conservative: 26  
Best Local Similarity: 38.75% Mismatches: 46  
Query Match: 21.98% Indels: 26  
DB: 2 Gaps: 5

US-10-728-051-2 (1-717) x A33090 (1-153)

Qy	2	CTCACCATACTAGTAGCCCTC---GCCCTTTTCTCTCTCGCTGCCACGCGATCTGCAGG 58
		:     :     :     :     :     :     :     :     :     :
Db	4	LeuThrlleuIleAlaLeuValAlaAlaLeuValLeuValValHisThrSerAlaPhe 23
		:     :     :     :     :     :     :     :     :
Qy	59	CAGCAGTGGGAACCTCCCAAGAGACACAGAGATGCCAGAGCCAGCTCGAGAGGGCGCAACTG 118
		:     :     :     :     :     :     :     :     :
Db	24	Gln-----SerSerLysGlnSerCysLysArgGlnLeuGlnGlnValAsnLeu 39
		:     :     :     :     :     :     :     :     :

```
QY 119 AGCCCTGCGAGCAATCTCATGCAAGATCAACGTCACGAGGATTCATATGAACGG 178
Db 40 ArgHisCysGluAsnHisIleAlaGlnArgIleGlnGlnGlnGlnGluGluAsp 59
QY 179 GACCCGTACAGC-----CCTAGTCAGGATCCG 205
Db 60 HisAlaLeuLysLeuArgGlyIleLysHisValIleLeuArgHisArgSerGlnGlu 79
QY 206 TACAGCCCTAGTCATATGATCGAGAGGGCGCTGCTCTCAGACCAAGAGAGTGT 265
Db 80 TyrSerGlu-----GluSerGluGluLeuAspGlnCys 90
QY 266 TGCATAGCTGAGCTGAACGAGTTTGAGAACCAACAAAGTGTGCGAGGCGATTGCAACAG 325
Db 91 CysGluGlnLeuAsnGluLeu---AsnSerGlnArgCysGlnCysArgAlaLeuGlnGln 109
QY 326 ATCATGAGAACGAGAGCGATAGTTGTCAGGGAGGCAACAGAGCAACAGTTCAAGAGG 385
Db 110 IleTyrGluSerGlnSerGlnGlnCysGluGlySerGlnGlnGlnGlnGlnGluGln 129
QY 386 GAGCTCAGGAACCTTGCTCAACAGTGTGGCGCTTAGGGCACCAAGCGTTGCGACTTGGAC 445
Db 130 GluLeuGlnLysLeuProArgThrCysGlyPheGlyProLeuArgArgCysAspValAsn 149

RESULT 2
T05710
2S albumin precursor - soybean
N/Alternate names: aspartic acid-rich peptide
C/Species: Glycine max (soybean)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: T05710; A28485
R/Wang, J.; Pichersky, E.
Plant Physiol. 114, 1567, 1997
A/Title: Nucleotide Sequence of S-Adenosyl-L-Methionine:Caffeic Acid 3-O-Methyltransferase
A/Reference number: Z15424
A/Accession: T05710
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-158 <WAN>
A/Cross-references: UNIPROT:P19594; EMBL:AF005030; NID:G2305019; PID:G2305020
A/Experimental source: cultivar Hodgson 78; cotyledon
R/Odani, S.; Koide, T.; Ono, T.
J. Biol. Chem. 262, 10502-10505, 1987
A/Title: Amino acid sequence of a soybean (Glycine max) seed polypeptide having a poly(U
A/Reference number: A28485; MUID:87280104; PMID:3611081
A/Accession: A28485
A/Molecule type: protein
A/Residues: 22-64 <ODA>
C/Superfamily: soybean 2S albumin
C/Keywords: glycoprotein
F/1-21/Domain: signal sequence #status predicted <SIG>
F/22-64/Product: 2S albumin small chain #status predicted <MAT>
F/54-56/Region: cell attachment (R-G-D) motif
F/39/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 7,86e-15 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.49% Conservative: 31
Best Local Similarity: 36.36% Mismatches: 55
Query Match: 19.21% Indels: 12
DB: 1 Gaps: 6

US-10-728-051-2 (1-717) x T05710 (1-158)

QY 5 ACCATAGTAGCCCTCGCCCTTTCTCCTCGTCCCGCCAGCATCTCGGAGGACGAG 64
Db 5 ThrIleLeuLeuIleSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGGAACCTCAAGGAGACAGATGCGCAGAGCGAGCTCGAGAGGGCGACCTGAGGCC 124
Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42
```

```
QY 125 TCGAGCAACATCTCATGCAAGATCCAA-----CGTGACGAGGATTCATATGAACGG 178
Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAsp 62
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGGCT 238
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgArgAsnGlu 82
QY 239 GGATCTCTCTCAG-----CACCAGAGAGGTTGCAATGAGCTGAAC 280
Db 83 GlyLysAspGluAspGluGluGluGluGluHisMetGlnLysCysCysThrGluMetSer 102
QY 281 GAGTTTGAGAACCAACCAAGGTGTCATGTCGAGGCGATTGCAACAGATCATGGAACACAG 340
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 AGCGATAGTTGTCAGGGAGGCAACAGAGCAACAGTTCACAGAGGAGCTCAGGAACCTTG 400
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysLysMetGluLysGluLeuIleAsnLeu 140
QY 401 CCTCAACAGTGTGGCGCTTAGGGCACCAAGCGTTGCGACTTG 442
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 3
B23617
conglutin delta-2 large chain - narrow-leaved blue lupine
C/Species: Lupinus angustifolius (narrow-leaved blue lupine)
C/Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
C/Accession: B23617
R/Lilley, G.G.; Inglis, A.S.
FEBS Lett. 195, 235-241, 1986
A/Title: Amino acid sequence of conglutin delta, a sulfur-rich seed protein of Lupinus ar
A/Reference number: A91358
A/Accession: B23617
A/Molecule type: protein
A/Residues: 1-80 <HIL>
A/Cross-references: UNIPROT:P09931
C/Superfamily: soybean 2S albumin

Alignment Scores:
Pred. No.: 1.42e-09 Length: 80
Score: 187.50 Matches: 34
Percent Similarity: 73.13% Conservative: 15
Best Local Similarity: 50.75% Mismatches: 17
Query Match: 14.82% Indels: 1
DB: 2 Gaps: 1

US-10-728-051-2 (1-717) x B23617 (1-80)

QY 245 TCTCAGCAACAGAGAGGTTGTCATGATGAGCTGAACGAGTTTGAGAACCAACCAAGGTGC 304
Db 11 SerGluGluLeuAspGlnCysCysGluGlnLeuAsnGluLeu---AsnSerGlnArgCys 29
QY 305 ATGTGCGAGGCAATTCACAGATCATGAGAACCGAGAGCGATGAGTTGCGAGGAGGCA 364
Db 30 GlnCysArgAlaLeuGlnIleTyrGluSerGlnSerGluGlnCysGluGlyArgGln 49
QY 365 CAGGAGCAACAGTTCAAGAGGAGGCTCAGGAACCTTCCTCAACAGTGGCGCTTAGGGCA 424
Db 50 GlnGluGlnGlnLeuGluGlyGluLeuGluLysLeuProArgIleCysGlyPheGlyPro 69
QY 425 CCACAGCGTTGCGACTTGGAC 445
Db 70 LeuArgArgCysAsnIleAsn 76

RESULT 4
S38887
2S albumin - common sunflower (fragment)
C/Species: Helianthus annuus (common sunflower)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: S38887
```

Alignment Scores:	2.97e-08	141
Pred. No.:		
Score:	173.50	52
Percent Similarity:	49.37%	Conservative: 26
Best Local Similarity:	32.91%	Mismatches: 47
Query Match:	13.72%	Indels: 33
DB:	2	Gaps: 10

US-10-728-051-2 (i-717) x T10257 (1-141)

Qy	2	CTCACCATACTAGTAGTACGGCTC---GCCCTTTTCTCTCGCTGCCTCCACGCGATCTCGCAGG	58
Db	4	LeuThrSerIleAlaLeuPheAlaValAlaLeuLeuValAlaAspAlaIatyrAlaIatyr	23
Qy	59	CAG-----CAGTGGGAACTCCAAGG---GACAGAGAATGC	91
Db	24	ArgThrThrIleThrValGluValGluLeuAsnArgGlnGlyArgGluArgCys	43
Qy	92	CAGACGACGCTCGAGAGCGGAACTGAGGCCCTGCGAGCAACATCTCATGCAGAGAATC	151
Db	44	ArgGlnMetSerAlaArgGluGluLeuArgSerCysGluGlnItyrLeuArgGlnGln---	62
Qy	152	CAACGTGACGAGGATTATATACAGCGGACCGGTACAGCCCTAGT-----CAGGATCCG	205
Db	63	-----SerArgAspValLeuGlnMetArgGlyIleGluAsnPro	75
Qy	206	TACAGCCCTTAGTCTCATATGATCGGAGAGCGCCTCGATCCTCTCAGCACCAAGAGAGGTGT	265
Db	76	Ttp-----ArgArgGluGlyGlySerPhe-----AspGluCys	86
Qy	266	TGCATGACGCTGAACGAGTTTGAGNACAACAAGAGTTCATGTGCGAGGCGATTTGCAACAG	325
Db	87	CysArgGluLeuLysAsnValAsp-----GluGlyCysArgCysAspMetLeuGluGlu	104
Qy	326	ATCATGTGGAAACACAGAGCGATAGTGTGAGGGGAGGCAACAGAGACCAACAGTTTCAAGAGG	385
Db	105	IleAlaArgGluGluGlnArgGlnAlaArgly---GlnGluGlyArgGlnMetLeuGln	123
Qy	386	GAGCTCAGGAACCTTCCTCAACAGTTCGGGCTTTAGGGCAACACAGCGTTGGGAC	439
Db	124	LysAlaArgAsnLeuProSerMetCysGlyIleArg---ProGlnArgCysAsp	140

## RESULT 6

2S albumin - Brazil nut  
C:Species: Bertholletia excelsa (Brazil nut)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 21-Jul-2000  
C:Accession: S14947  
P:Gander, E.S.; Holmstroem, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro  
Plant Mol. Biol. 16, 437-448, 1991  
A:Title: Isolation, characterization and expression of a gene coding for a  
A:Reference number: S14946; MUID:913370890; PMID:1840683

A;Accession: J14747  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-154 <GAN>  
A;Cross-references: EMBL:X54491; NID:g17712; PIDD:CAA38363.1; PID:g17713  
C;Genetics:  
A;Introns: 67/3  
C;Superfamily: wheat alpha-amylase inhibitor

Alignment Scores:  
Pred. No.: 2.07e-06 Length: 154

Alignment Scores:	2.07e-06	Length:	154
Pred. No.:			

Score: 154.00 Matches: 45  
 Percent Similarity: 45.34% Conservative: 28  
 Best Local Similarity: 27.95% Mismatches: 52  
 Query Match: 12.17% Indels: 36  
 DB: 2 Gaps: 10

US-10-728-051-2 (1-717) x S14947 (1-154)

QY 8 ATACTAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCGATCTGCG----- 55  
 Db 7 ValAlaAlaLeuLeuAlaLeuValLeuGluGlnAlaThrAlaPheArgThrThr 26  
 QY 56 -----AGCAGCAGTGGAAATCCCAAGCA-----GACAGAAGA 88  
 Db 27 ValThrThrThrLeuGluGluGluGluGluGluGluGluGluGluGluGlu 46  
 QY 89 TGCAGAGCCAGCTCGAGAGG---GCGAAGCTGAGGCCCTCGAGCAACATCTCATGCG 145  
 Db 47 CysArgGluGluMetGluArgGluGluGluGluGluGluGluGluGluGlu 66  
 QY 146 AAGATCCAAAGCTGAGGAGGATTCATATCAACGGGACCCGTACAGCCCTAGTCAGGATCCG 205  
 Db 67 GlnMet-----GluGluSerProTyrGln----- 74  
 QY 206 TACAGCCCTAGTCATATGATCGAGAGGCGCTGATCTCTCAGCAACAGAGAGGTGT 265  
 Db 75 ---AsnProArgProLeuArgArg-----GlyGluGluProHisLeuAspGluCys 90  
 QY 266 TGCAATGAGTGAACGAGTTTGAGAACACCAAGGTGCGATGTCGAGGCGATTCGCAACAG 325  
 Db 91 CysGluGluGluGluArgMetAsp-----GluMetCysArgCysGluGluGluArgMet 108  
 QY 326 ATCATGAGAACCAAGAGCGAT-----AGGTTGAGGGGAGGCAACAGAGCAACAGTTC 379  
 Db 109 MetLeuArgArgGlnArgGluGluGluGluGluGluGluGluGluGluGlu 127  
 QY 380 AAGAGGAGCTCAGGAAGTTCCTCAACAGTGGCGCTTAGGGCACCAGAGCGTTGCGAC 439  
 Db 128 MetArgGlyAlaGluAsnLeuLeuSerArgCysAsnLeu---SerProGlnArgCysPro 146  
 QY 440 TTG 442  
 Db 147 Met 147

# RESULT 7

S01062  
 2S seed storage protein precursor (clone HaG5) - common sunflower  
 N;Alternate names: 2S albumin storage protein  
 C;Species: Helianthus annuus (common sunflower)  
 C;Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004  
 C;Accession: S01062  
 R;Allen, R.D.; Cohen, E.A.; Vonder Haar, R.A.; Adams, C.A.; Ma, D.P.; Nessler, C.L.; Tho  
 Mol. Gen. Genet. 210, 211-218, 1987  
 A;Title: Sequence and expression of a gene encoding an albumin storage protein in sunfl  
 A;Reference number: S01062; MUID:88142538; PMID:2830455  
 A;Accession: S01062  
 A;Molecule type: DNA  
 A;Residues: 1-295 <ALL>  
 A;Cross-references: UNIPROT:P15461; EMBL:X06410; NID:g18799; PIDN:CAA29699.1; PID:g18800  
 A;Note: part of this sequence, including the amino end of the mature protein, was confir  
 C;Genetics:  
 A;Introns: 192/2  
 C;Superfamily: gliadin  
 F;1-20/Domain: signal sequence #status predicted <SIG>  
 F;162-295/Product: 2S seed storage protein #status experimental <NAT>

Alignment Scores:  
 Pred. No.: 3 93e-06 Length: 295  
 Score: 151.00 Matches: 47  
 Percent Similarity: 40.39% Conservative: 35  
 Best Local Similarity: 23.15% Mismatches: 65  
 Query Match: 11.94% Indels: 56  
 DB: 2 Gaps: 8

US-10-728-051-2 (1-717) x S01062 (1-295)

QY 8 ATACTAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCGATCTGCG----- 55  
 Db 5 IleValIleAlaLeuAlaPheAlaLeuValAlaPheAlaThrAlaThrThrThr 24  
 QY 56 -----AGCAGCAGTGGGAATCCCAAGGAGACAGAGATGCCAGGCCAG 100  
 Db 25 IleThrThrThrIleGluAspGluAsnProIleSerGlyGlnArgGlnValSerGlnArg 44  
 QY 101 CTCGAGAGCGGCAACCTCGAGGCCCTCGAGCAACATCTCTCATG-----CAGAAGATCCAA 154  
 Db 45 IleGlnGlyGlnArgLeuAsnGlnCysArgMetPheLeuGlnGlnGlyGlnAsnIlePro 64  
 QY 155 CGTGACGACGATTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCT 214  
 Db 65 ArgGluPheAspAsn----- 69  
 QY 215 AGTCCATATGATCGGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGTTGCAATGAG 274  
 Db 70 -----ProGlnMetGlyArgGlnGlnGluGlnGlnLeuGlnGlnCysCysGlnGlu 86  
 QY 275 CTCACGAGTTTGAGAACCAACCAAGAGTGCATGTGCGAGGCATTGCAACAGATCATGCGAG 334  
 Db 87 LeuGlnAsnIleGluGly-----GlnCysGlnCysGluAlaValysGlnValPheArg 104  
 QY 335 AACGAGCGATGAGTTGCGAG-----GGGAGG 361  
 Db 105 GluAlaGlnGlnGlnValGlnGlnGlnGlnGlnGlnGlnLeuValProPheArgGlySer 124  
 QY 362 CAACGAGGCAACAGTTCAAGAGGAGGAGCTCAGGAACCTTGCCTCAACAGTGGCGCTTAGG 421  
 Db 125 GlnGlnThrGlnGlnLeuLysGlnLysAlaGlnIleLeuProAsnValCysAsnLeuGln 144  
 QY 422 GCACACACGCGTTGCGAGTTG-----GACGTCGAA 451  
 Db 145 Ser--ArgArgCysGluIleGlyThrIleThrThrThrValThrGluSerAsnIleAsp 163  
 QY 452 AGTGGCGGACAGACAGATACTAAACACTATCTCAAAAAAGAAAAAGAAAAAGAA 511  
 Db 164 IleProPheArgAspArg-----ProPheGlyThrGlySerGlnGlnCysArgGlu 180  
 QY 512 AATAGCTTA 520  
 Db 181 ThrGluIle 183

# RESULT 8

A29802  
 napin precursor (gNa) - rape  
 N;Alternate names: 1.7 S seed storage protein  
 C;Species: Brassica napus (rape)  
 C;Date: 20-Jun-1989 #sequence\_revision 20-Jun-1989 #text\_change 09-Jul-2004  
 C;Accession: A29802  
 R;Scofield, S.R.; Crouch, M.L.  
 J. Biol. Chem. 262, 12202-12208, 1987  
 A;Title: Nucleotide sequence of a member of the napin storage protein family from Brassic  
 A;Reference number: A29802; MUID:87308225; PMID:3040733  
 A;Accession: A29802  
 A;Molecule type: DNA  
 A;Residues: 1-186 <SCO>  
 A;Cross-references: UNIPROT:P09893; GB:J02782; NID:g167178; PIDN:AAA33007.1; PID:g167179  
 C;Genetics:  
 A;Gene: gNa  
 C;Superfamily: wheat alpha-amylase inhibitor

Alignment Scores:  
 Pred. No.: 9 49e-06 Length: 186  
 Score: 147.00 Matches: 47  
 Percent Similarity: 40.98% Conservative: 28  
 Best Local Similarity: 25.68% Mismatches: 52  
 Query Match: 11.62% Indels: 56  
 DB: 2 Gaps: 9

```
US-10-728-051-2 (1-717) x A29802 (1-186)
QY 8 ATACTAGTACCTCGCCCTTTTCCTCTCGCTGCCCGCATCTGCG---AGGCAGCAG 64
   ::::|||||
Db 8 ValSerAlaThrLeuAlaLeuPhePheLeuLeuThrAsnAlaSerValTyrArgThrVal 27
   ::::|||||
QY 65 TGGGAAGTCCAGAGAGAC-----AGAAAGATGC 91
   ::::|||||
Db 28 ValGluValAspGluAspAspAlaThrAsnProAlaGlyProPheArgGlieProLysCys 47
   ::::|||||
QY 92 CAGAGCCAGCTCGAGAGGGCG---AACCTGAGGCCCTGCGCAGCAACATCTCATGCGAGAAG 148
   ::::|||||
Db 48 ArgLysGluPheGlnGlnAlaGlnHisLeuArgAlaCysGlnGlnTrpLeuHisLysGln 67
   ::::|||||
QY 149 ATCAACAGCTGAC-----160
   ::::
Db 68 AlaMetGlnProGlyGlyGlySerGlyProSerTrpThrLeuAspGlyGluPheAspPhe 87
   ::::
QY 161 GAGATTATATGAACGGGACCGGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCCA 220
   ::::|||||
Db 88 GluAspAspValGluAsnGlnGlnGlnGlyProGlnGlnArgPro-----ProProPro 105
   ::::|||||
QY 221 TATGATCGGAGAGCGCTGATCTCTCAGCACCAGAGAGGTGTTCATGAGCTGAAC 280
   ::::|||||
Db 106 -----GlnGlnCysCysAsnGluLeuHis 113
   ::::|||||
QY 281 GAGTTTGAGAACCAACAAAGGTGCATGTGCGAGGCATTG-----319
   ::::|||||
Db 114 GlnGlu-----ProLeuCysValCysProThrLeuLysGlyAlaSerLysAlaVal 131
   ::::|||||
QY 320 ---CAACAGATCATGGAGAACCGAGAGCATAGTTGCGGGAGGCGCAACAGGACGACAG 376
   ::::|||||
Db 132 ArgGlnGlnValArgGlnGlnGlnGlnGlnMetGlnGlnGlnMetGlnGlnVal 151
   ::::|||||
QY 377 TTCAAGAGGAGCTCAGG-----AACTGCCTCAACAGTGGCGCCTTAGGCGACCA 427
   ::::|||||
Db 152 IleSerArgValTyrGlnThrAlaThrHisLeuProArgValCysAsnIleLeuGlnVal 171
   ::::|||||
QY 428 CAGCGTTGC 436
   ::::
Db 172 SerIleCys 174

RESULT 9
NWMTU
2S albumin 1 precursor - Arabidopsis thaliana
N;Alternate names: seed storage protein AT2S1
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: JAO161; PS0282; S34676; T06044
R;Krebbbers, E.; Herdies, L.; de Clercq, A.; Seurinck, J.; Van Damme, J.; Se
Plant Physiol. 87, 859-866, 1988
A;Title: Determination of the processing sites of an Arabidopsis 2S albumin and character
A;Reference number: JAO161
A;Accession: JAO161
A;Molecule type: DNA
A;Residues: 1-164 <KR2>
A;Cross-references: UNIPROT:P15457; GB:M22032; NID:g166609; PIDN:AAA32743.1; PID:g166614
A;Accession: PS0282
A;Molecule type: protein
R;Conceicao, A.D.S.; Krebbbers, E.
submitted to the EMBL Data Library, July 1993
A;Description: Tentative title: a cotyledon regulatory region is responsible for the dif
A;Reference number: S34674
A;Accession: S34676
A;Molecule type: DNA
A;Residues: 1-164 <CON>
A;Cross-references: EMBL:Z24745; NID:g395203; PIDN:CAA80870.1; PID:g395204
R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15484
A;Accession: T06044
```

```
A;Molecule type: DNA
A;Residues: 1-164 <BEV>
A;Cross-references: EMBL:AL035680; GSPDB:GN00062; ATSP:T24A18.90
A;Experimental source: cultivar Columbia; BAC clone T24A18
C;Genetics:
A;Gene: T24A18.90
A;Map position: 4
C;Superfamily: wheat alpha-amylase inhibitor
C;Keywords: seed; storage protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-164/Product: 2S albumin 1 proprotein #status predicted <AT2>
F;38-73/Product: 2S albumin 1 small chain #status experimental <SMC>
F;84-162/Product: 2S albumin 1 large chain #status experimental <LAC>

Alignment Scores:
Pred. No.: 1.06e-05 Length: 164
Score: 146.50 Matches: 46
Percent Similarity: 46.47% Conservative: 33
Best Local Similarity: 27.06% Mismatches: 52
Query Match: 11.58% Indels: 39
DB: 1 Gaps: 9

US-10-728-051-2 (1-717) x NWMTU (1-164)
QY 8 ATACTAGTACCTCGCCCTTTTCCTCTCGCTGCCCGCATCTGCG---AGGCAGCAG 64
   ::::|||||
Db 8 ValCysAlaAlaLeuAlaLeuCysPheLeuLeuThrAsnAlaSerIleTyrArgThrVal 27
   ::::|||||
QY 65 TGGGAAGTCCAGAGAGAC-----AGAAAGATGCCAG 94
   ::::|||||
Db 28 ValGluPheGlnLysAspAlaThrAsnProIleGlyProLysMetArgLysCysArg 47
   ::::|||||
QY 95 AGCAGCTCGAGAGG---GCGAACCTGAGCCCTGCGAGCAACATCTCATGCAAGATC 151
   ::::|||||
Db 48 LysGluPheGlnLysGlnGlnHisLeuArgAlaCysGlnGlnLeuMetLeuGlnAla 67
   ::::|||||
QY 152 CAACGT-----CAGCAGGATTATATGACGGGAGCCGCTACAGCCCTAGTCAGGAT 202
   ::::|||||
Db 68 ArgGlnGlyArgSerGluPheAspPheGluAspMetGluAsnProGln-----85
   ::::|||||
QY 203 CCGTACAGCCCTAGTCCATATGATGCGAGAGCGCGCTCTCTCAGACCAA-----256
   ::::|||||
Db 86 -----GlyGlnGlnGlnGlnGlnLeu 93
   ::::|||||
QY 257 ---GAGAGGTGTTGCAATGAGCTGAACGAGTTTGAACAACCAAGGTGTCATGTGCGAG 313
   ::::|||||
Db 94 PheGlnGlnCysCysAsnGluLeuArgGlnGluGluProAsp-----CysValCysPro 111
   ::::|||||
QY 314 GCATTGCAACAGATCATGGAGAACCGAGCGCATAGTTGCGAGGAGGCAACAGGACAA 373
   ::::|||||
Db 112 ThrLeuLysGlnAlaAla-----LysAlaValArgLeuGlnGlnHisGlnProMet 129
   ::::|||||
QY 374 CAGTTCAAGAGG-----GAGCTCAGGAACCTGCTCAACAGTGGCGGCTTAGGGCA 424
   ::::|||||
Db 130 GlnValArgLysIleTyrGlnThrAlaLysHisLeuProAsnValCysAspIleProGln 149
   ::::|||||
QY 425 CCACAGCGTTGCGACTTGGACGTGCAAGT 454
   ::::|||||
Db 150 ValAspValCysPropheAsnIleProSer 159

RESULT 10
S14946
2S seed storage protein large chain - Brazil nut
N;Alternate names: albumin 2S precursor
C;Species: Bertholletia excelsa (Brazil nut)
C;Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: S14946; S06252; S21640; B5802
R;Gander, E.S.; Holmstrom, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro, M.; Gross
Plant Mol. Biol. 16, 437-448, 1991
A;Title: Isolation, characterization and expression of a gene coding for a 2S albumin fr
A;Reference number: S14946; MUID:91370890; PMID:1840683
A;Accession: S14946
A;Status: preliminary
```

Db 70 Pro-----ArgArg-----GlyMetGluProHisMetSerGluCysCysGluGlnLeu 85  
 QY 278 AACGAGTTTGAGAAACAACAAAGGTGCATGTCGAGGCGATTGCCAA---CAGATCATGGAG 334  
 Db 86 GluGlyMetAsp-----GluSerCysArgCysGluGlyLeuArgMetMetMetMetArg 108  
 QY 335 AACCAGAGCGATAGTTGTCGAGGGAGG---CAACAGGAGCAACACAGTTCACAGAGGAGCTC 391  
 Db 104 MetGlnGlnGluMetGlnProArgGlyGluGlnMetArgMetMetArgLeuAla 128  
 QY 392 AGGAACCTTCCTCAACAGTGCCTTAGCGCCCTTAGGCGACCAACAGCGTTTCGCACTTG 442  
 Db 124 GluAsnIleProSerArgCysAsnLeu---SerProMetArgCysProMet 139

RESULT 11  
 JC5379  
 mabinlin II precursor - Yunnan caper  
 C:Species: Capparis maaikai (Yunnan caper)  
 C:Date: 02-Jun-1997 #sequence\_revision 12-Sep-1997 #text\_change 09-Jul-2004  
 C:Accession: JC5379; PC4316  
 R:Nirasawa, S.; Masuda, Y.; Nakaya, K.; Kurihara, Y.  
 Gene 181, 225-227, 1996  
 A:Title: Cloning and sequencing of a cDNA encoding a heat-stable sweet protein  
 A:Reference number: JC5379; MUID:97128796; PMID:8973336  
 A:Accession: JC5379  
 A:Molecule type: mRNA  
 A:Residues: 1-155 <NIR>  
 A:Cross-references: UNIPROT:P30233; DDBJ:D83997; NID:G1817545; PIDN:BAAL2204.1  
 A:Accession: PC4316  
 A:Molecule type: protein  
 A:Residues: 36-41;149-154 <NIR2>  
 A:Experimental source: seed  
 C:Superfamily: wheat alpha-amylase inhibitor  
 C:Keywords: sweet taste  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-35/Domain: amino-terminal propeptide #status predicted <PRO>  
 F:36-68,83-154/Product: mabinlin II #status experimental <MAT>  
 F:36-68/Domain: mabinlin II A chain #status experimental <CHA>  
 F:69-83/Domain: mabinlin linker peptide #status predicted <LNK>  
 F:83-154/Domain: mabinlin II B chain #status experimental <CHB>

Alignment Scores:	
Pred. No.:	5,44e-05
Score:	139.00
Percent Similarity:	45.51%
Best Local Similarity:	28.85%
Query Match:	10.99%
DB:	2
	Gaps: 9

US-10-728-051-2 (1-717) x JC5379 (1-155)

QY 8 ATACTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCATCTCGAGG----- 58  
 Db 7 LeuPheAlaThrLeuAlaLeuPheValLeuAlaAsnAlaSerIleGlnThrVal 26  
 QY 59 -----CAGAGTGGGAACCTCCAGGAGACAGAAGATCC-----ArgCys 40  
 Db 27 IleGluValAspGluGluAspAsnGlnLeuTrp----- 40  
 QY 92 CAGAGCCAG---CTCGAGAGGGCGCAACTCAGAGCCCTCGGAGCAACATCTCATGCGAGAG 148  
 Db 41 GlnArgGlnPheLeuGlnHisGlnArgLeuArgAlaCysGlnArgPheIleHisArg 60  
 QY 149 ATCCAA-----CGTACGAGGATTTCATATGAACGGGACCCGTACAGCCCTAGTCAGGAT 202  
 Db 61 AlaGlnPheGlyGlyGlnProAspGluLeuGluAspGluValGluAspAsnAsp 80  
 QY 203 CGGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTGGATCTCTCTCAGCACCAAGAGAGG 262  
 Db 81 GluAsnGlnPro-----ArgArgProIa-----LeuArgGln 91  
 QY 263 TGTTCGAATGAGCTGAACGAGTTTGAGAAACCAAAAGTGATGTGC----- 310

Db 92 CysCysAsnGlnLeuArgGlnValAsp-----ArgProCysValCysProValLeuArg 109  
QY 311 GAGGCATTCACACAGATCATGAGAACACAGAGGATAGGTTCAGGGGAGGACACAGAG 370  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 110 GlnAlaAlaGlnGlnValLeuArgGlnLe-----lleGlnGlyProGlnGlnLeu 127  
QY 371 CAACAGTTCAAGAGGAGCTCAGAACTTCGCTCAACAGTGGCGGCTT 418  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 128 ArgArgLeuPheAspAlaAlaArgAsnLeuProAsnIleCysAsnIle 143  
RZCS  
2S seed storage protein precursor - castor bean  
N;Alternate names: 2S albumin precursor  
C;Species: Ricinus communis (castor bean)  
C;Date: 14-Nov-1983 #sequence revision 08-Feb-1996 #text change 09-Jul-2004  
C;Accession: S11499; S11500; S11501; S11502; S27221; A01328; S27222  
R;Irwin, S.D.; Lord, J.M., 1990  
Nucleic Acids Res. 18, 5890, 1990  
A;Title: Nucleotide sequence of a Ricinus communis 2S albumin precursor gene.  
A;Reference number: S11499; MUID:91016940; PMID:2216785  
A;Accession: S11499  
A;Molecule type: DNA  
A;Residues: 1-258 <IR>  
A;Cross-references: UNIPROT:P01089; EMBL:X54158; NID:921067; PID:CAA38097.1; PID:921068  
A;Note: the authors translated the codon CTC for residue 14 as Phe, CCA for residue 74 as  
R;Irwin, S.D.; Keen, J.N.; Findlay, J.B.C.; Lord, J.M.  
Mol. Gen. Genet. 222, 400-408, 1990  
A;Title: The Ricinus communis 2S albumin precursor; a single preproprotein may be processed  
A;Reference number: S11500; MUID:91109729; PMID:2274038  
A;Accession: S11500  
A;Molecule type: mRNA  
A;Residues: 1-13,'F',15-73,'T',75-258 <IR2>  
A;Experimental source: clone 14g4  
A;Accession: S11501  
A;Molecule type: mRNA  
A;Residues: 'N',4,'LS',7-13,'F',15-21 <IRF>  
A;Experimental source: clone 10a12  
A;Accession: S27221  
A;Molecule type: protein  
A;Residues: 'X',37,'X',39-45,'X',158-161,'X',163-174,'X', <IRA>  
R;Sharief, F.S.; Li, S.S.L.  
J. Biol. Chem. 257, 14753-14759, 1982  
A;Title: Amino acid sequence of small and large subunits of seed storage protein from Ricinus communis  
A;Reference number: A92357; MUID:83082772; PMID:7174664  
A;Accession: A01328  
A;Molecule type: protein  
A;Residues: 157-190:194-221,'Q',223-225,230-233,'N',235-254,'Q',256-258 <SHA>  
A;Note: 230-Ser was also found  
A;Note: there is considerable similarity between residues 181-231 of this protein and residues 181-231 of the protein from Ricinus communis  
B;Odan, S.; Koide, T.; Ono, T.; Ohnishi, K.  
Biochem. J. 213, 543-545, 1983  
A;Title: Structural relationship between barley (Hordeum vulgare) trypsin inhibitor and soybean trypsin inhibitor  
A;Reference number: A90322; MUID:83308577; PMID:6615448  
A;Contents: annotation  
A;Note: this protein is homologous with trypsin inhibitor from barley  
C;Complex: consists of two chains linked by two disulfide bonds involving Cys-162 and Cys-163  
C;Superfamily: 2S seed storage protein precursor  
C;Keywords: pyroglutamic acid; seed; storage protein  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-35/Domain: propeptide #status predicted <PRO>  
F;36-72/Product: probable 2S seed storage protein small chain 2 #status experimental <SC>  
F;87-156/Product: probable 2S seed storage protein large 2 #status predicted <LCHI>  
F;157-190/Product: 2S seed storage protein small chain #status experimental <SML>  
F;194-258/Product: 2S seed storage protein large chain #status experimental <LRG>  
F;194/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experime

Alignment Scores:

Pred. No.: 5.38e-05 Length: 258

Score: 139.00 Matches: 40  
Percent Similarity: 48.03% Conservative: 33  
Best Local Similarity: 26.32% Mismatches: 51  
Query Match: 10.99% Indels: 28  
DB: 1 Gaps: 7

US-10-728-051-2 (1-717) x RZCS (1-258)

QY 11 CTAGTAGCCCTCGCCCTCTTCTCTCTCGTCCCGCCGATCTCGAGGCGACGAGTGG--- 67

Db 10 LeuValSerValLeuLeuPheIleIleAlaAsnAlaSerPheAlaTyrArgThrThrIle 29

QY 68 -----GAATCCCAAGAGACAGAGA-----AGATGC 91

Db 30 ThrThrIleGluIleAspGluSerLysGlyGluArgGluGlySerSerSerGlnGlnCys 49

QY 92 CAGAGCCAGCTCCAGAGGGCGACCTGAGCCCTCGAGCAACATCTCATGCAGAAGATC 151

Db 50 ArgGlnGluValGlnArgLysAspLeuSerSerCysGluArgTyrLeuArgGlnSerSer 69

QY 152 CAACGTGACGAGGATTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGC 211

Db 70 SerArg-----ArgSerProGlyGluGluValLeuArg 80

QY 212 CCTAGTCCATATGATCGGAGAGCGGTGATCTCTCAGACCACAGAGAGGTGTGCAAT 271

Db 81 ---MetProGlyAspGluAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 99

QY 272 GAGCTCAACGAGTTTCAGAACACAAAGTGTGATGCGAGGCGATTCGCAACAGATCATG 331

Db 100 GlnValLysGlnValArgAsp-----GluCysGlnCysGluAlaIleTyrIleAla 117

QY 332 GAGAAC-----CAGAGCGATAGTTTCAGGAGGAGGACCAACAGAGAGGTTCAGAGG 385

Db 118 GluAspGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 137

QY 386 GAGCTCAGAACTTGGCTCAACAGTGGCGGCTTAGG 421

Db 138 ---AlaGlyGluIleValSerSerCysGlyValArg 148

RESULT 13

T09252

seed storage protein EMB25 - white spruce

C;Species: Picea glauca (white spruce)

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C;Accession: T09252

R;Dong, J.Z.; Dunstan, D.I.

submitted to the EMBL Data Library, June 1996

A;Description: Gene expression during somatic embryogenesis.

A;Reference number: Z16588

A;Accession: T09252

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-165 &lt;DON&gt;

A;Cross-references: UNIPROT:Q40850; EMBL:L47745; NID:gl350513; PID:gl350514

C;Genetics:

A;Gene: EMB25

Alignment Scores:

Pred. No.: 0.000224 Length: 165

Score: 132.50 Matches: 47

Percent Similarity: 38.42% Conservative: 21

Best Local Similarity: 26.55% Mismatches: 44

Query Match: 10.47% Indels: 65

DB: 2 Gaps: 10

US-10-728-051-2 (1-717) x T09252 (1-165)

QY 8 ATACTAGTACCCCTCGCCCTTTTCTCTCTCGCT-----GCC 43

Db 17 ValSerLeuGlyMetAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 36

QY 44 CACGCATCTCGAGGCGAGCATGGGAACTCCAAGAGACAGAGA---TGCCAGAGCCAG 100



```
Db 37 HisGluAspGlyLeuTyrGlyGluValGlnGlnArgSerCysGluGlnGln 56
QY 101 CTCGAGAGGCGCAACTGAGGCCCTCGGAGCAACATCTCATGCAGAAATCCACGTGAC 160
Db 57 -----ArgLeuSerSerCysArgGluTyrLeu-----GluArgPro 68
QY 161 GAGGATTTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCA 220
Db 69 ArgAsp----- 70
QY 221 TATGATCGAGAGGCGCTGGATCTCTCAGCACCAAGAGAGGTGTTCATGAGCTGAAC 280
Db 71 -----GlnProSerGluArgCysGluGluGln 81
QY 281 GAGTTTGAGAACCAAGGTGCATGTCGAGGCAATTCGAACAGATCATGGAAAC--- 337
Db 82 ArgMet-----SerProGlnCysArgCysGlnAlaIleGlnArgThrLeuGluAspVal 99
QY 338 -----CAGAGCGAT-----AGG 349
Db 100 PheMetAspSerAspSerGlnAspGlyAlaProLeuAsnGlnArgArgGlnArgArg 119
QY 350 TTGAGGCGGAGG---CAACAGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCCCTCAA 406
Db 120 GlyGlnGlyArgGlyMetGluGluGluValValArgAlaGluGluLeuProAsn 139
QY 407 CAGTGGCGCCCTAGG---GCACCAACAGCGTTGCGACCTTGAGCGTCGAAAGT 454
Db 140 ThrCysAsnValArgGlnSerProArgArgCysAspLeuGlnArgHisSer 156

RESULT 14
T08012
2S seed storage protein precursor - Douglas fir
C:Species: Pseudotsuga menziesii (Douglas fir)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C:Accession: T08012
R:Chatthai, M.; Misra, S.
Planta 206, 138-145, 1998
A:Title: Structure and expression of embryogenesis-specific cDNAs encoding 2S seed storage
A:Reference number: 216286; MUID:98381325; PMID:9715536
A:Accession: T08012
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-170 <CH>
A:Cross-references: UNIPROT:O64931; EMBL:AF029972; NID:G2988481; PIDN:AAC27000.1; PID:G2
C:Genetics:
A:Gene: 2S3
C:Keywords: seed; storage protein
```

```
Alignment Scores:
Pred. No.: 0.00048 Length: 170
Score: 129.00 Matches: 43
Percent Similarity: 36.11% Conservative: 22
Best Local Similarity: 23.89% Mismatches: 49
Query Match: 10.20% Indels: 66
DB: 2 Gaps: 8

US-10-728-051-2 (1-717) x T08012 (1-170)
QY 8 ATACTAGTACGCTCGCCCTTTTCTCTCGTGGTCCCGCCATCTCGAGGCGAGTGG 67
Db 17 ValSerLeuGlyValAlaLeuLeuLeuVal-----GlnTrp 29
QY 68 -----GAACCTCAAGGAGACAGAAAGATGCCAGAGC 97
Db 30 SerThrProAsnValAspAlaGlyAspAsnMetPheGlyGluAspValValGlnGln 49
QY 98 CAGCTCGAGAGGCGCAACCTGAGGCCCTCGAGCAACATCTCATGCAGAGATCCAAAGT 157
Db 50 GlnGlnArgArgGlySerCysAspPro-----GlnArgLeuSerSer 63
QY 158 GACGAGGATTTCATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGT 217
```

```
Db 64 CysArgAspTyrLeuGluArgArgArgGluGlnProSer----- 76
QY 218 CCATATGATCGGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTG 277
Db 77 -----GluArgCysCysAsnGluLeu 83
QY 278 AACGAGTTTGAGAACCAACAAAGGTGCATGTGCGAGGCAATTCGAACAGATCATGGAGAAC 337
Db 84 GluArgMet-----SerProGlnCysArgCysProAlaIleGlnGlnValPheAspGln 101
QY 338 CAGAGCGATAGGTG----- 352
Db 102 SerSerGluAspLeuSerMetValAspSerHisSerGlnAsnAlaIleGlnArgGln 121
QY 353 -----CAGGAGGAGCAACAG---GAGCAACAGTTCAAGAGGAGCTCAGGAAC 397
Db 122 ArgArgGluGluArgGlyArgGluGluAlaGluGluMetValGluArgAlaGlnArg 141
QY 398 TTGCCTCAACAGTGGCGCTTAGGGCACCAAGCGT---TGGACTTGGAGCTCGAAAGT 454
Db 142 LeuProAsnThrCysAsnValArgGlnProProArgHisCysAspIleGlnArgHisSer 161

RESULT 15
S18871
2S-like storage protein - white spruce
C:Species: Picea glauca (white spruce)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S18871
R:Newton, C.H.
submitted to the EMBL Data Library, November 1991
A:Reference number: S18871
A:Accession: S18871
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-172 <NEW>
A:Cross-references: UNIPROT:P26986; EMBL:X63193; NID:G20495; PID:G20496

Alignment Scores:
Pred. No.: 0.00056 Length: 172
Score: 128.00 Matches: 46
Percent Similarity: 37.43% Conservative: 21
Best Local Similarity: 25.70% Mismatches: 46
Query Match: 10.12% Indels: 66
DB: 2 Gaps: 9

US-10-728-051-2 (1-717) x S18871 (1-172)
QY 14 GTAGCCCTCGCCCTTTTCTCTCTC-----GTCGCCACGCA 49
Db 18 LeuSerValAlaLeuPheLeuLeuPheHisTrpGlyLeuProSerValAspGlyHisGlu 37
QY 50 TCTGGAGGAGCAGTGGGAACCTCAAGGAGACAGAGA---TCCAGAGCCAGCTCGAG 106
Db 38 AspAsnMetTyrGlyGluGluIleGlnGlnArgSerCysAspProGlnArgAsp 57
QY 107 AGGCGCAACCTGAGGCCCTCGAGCAACATCTCATGTGAGAGATCCAACTGACGAGAT 166
Db 58 ProGlnArgLeuSerSerCysArgAspTyrLeu----- 68
QY 167 TCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGAT 226
Db 69 -----GluArgArgGluGlnProSer----- 76
QY 227 CGGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTGAACGAGTTT 286
Db 77 -----GluArgCysCysGluGluLeuGlnArgMet 86
QY 287 GAGAACCAACAAAGGTGCATGTGCGAGGCAATTCGAACAGATCATGGAG----- 334
Db 87 -----SerProGlnCysArgCysGlnAlaIleGlnMetLeuAspGlnSerLeuSer 104
QY 335 -----AACGAGCGGATAGG 349
```

```
Db      105 TyrAspSerPheMetAspSerAspSerGlnGluAspAlaProLeuAsnGlnArgArgArg 124
QY      350 TTG-----CAGGGAGGCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTG 400
Db      125 ArgArgGluGlyArgGlyArgGluGluGluAlaMetGluArg---AlaAlaTyrLeu 143
QY      401 CCTCAACAGTGGCGCCTTAGG---GCACCACAGCGTTGCGACTTGGACGTCGAAAGT 454
Db      144 ProAsnThrCysAsnValArgGluProProArgArgCysAspIleGlnArgHisSer 162
```

Search completed: April 8, 2005, 05:30:20  
Job time : 36.3307 secs

**This Page Blank (uspto)**

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	842	66.6	179	2	Q647H0	Q647h0 atachis hyp
2	838	66.2	175	2	Q6PSU1	Q6psu1 atachis hyp
3	819	64.7	156	2	Q941R0	Q941r0 atachis hyp
4	815	64.4	172	2	Q8GV20	Q8gv20 atachis hyp
5	815	64.4	187	2	Q6FSU2	Q6fsu2 atachis hyp
6	808	63.9	169	2	Q7Y1C0	Q7y1c0 atachis hyp
7	792	62.6	166	2	Q84TU1	Q84tu1 atachis hyp
8	454.5	35.9	145	2	Q647G9	Q647g9 atachis hyp
9	443.5	35.1	144	2	Q8W251	Q8w251 atachis hyp
10	388.5	30.7	129	2	Q9SGG5	Q9sgg5 atachis hyp
11	300	23.7	160	2	Q9SQH1	Q9sqh1 atachis hyp
12	299	23.6	158	2	Q647G8	Q647g8 atachis hyp
13	278	22.0	153	2	Q99235	Q99235 lupinus ang
14	243.5	19.2	158	1	2SS_SOYBN	P19594 glycine max
15	241.5	19.1	155	2	Q9ZNZ4	Q9znz4 glycine max
16	187.5	14.8	80	1	CG2L_LUPAN	P09931 lupinus ang

```

QY 122 CCTGCGAGCAATCTCATGCGAGATCCAAAGTGCAGGATTCATATGAACGGGAC 181
Db 63 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 82
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCGGAGGCGCTGGA 241
Db 83 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 102
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACAGTTCAGAACCAACAAAGG 301
Db 103 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 122
QY 302 TGATGTGCGAGGATTCGAACAGATCATGAGAACACAGAGCATAGTTGCGAGGAGG 361
Db 123 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 142
QY 362 CACACGAGCAACAGTTCAGAGGGAGCTCAGGAATTCCTGCTCAACAGTCCGGCCCTAGG 421
Db 143 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 162
QY 422 GCACCCAGCGTTCGCACTTGGACGTGCAAAAGTGGCGGAGAGACAGATAC 472
Db 163 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 179

RESULT 2
Q6PSU1
ID Q6PSU1 PRELIMINARY; PRT; 175 AA.
AC Q6PSU1;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE Seed storage protein SSP2 (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Liao B., Li H., Lin X., Huang S.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY561854; AAT00599.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON TER 1
SQ SEQUENCE 175 AA; 20463 MW; B3FB5806B70733DE CRC64;

Alignment Scores:
Pred. No.: 5.77e-65 Length: 175
Score: 838.00 Matches: 156
Percent Similarity: 99.36% Conservative: 0
Best Local Similarity: 99.36% Mismatches: 1
Query Match: 66.25% Indels: 0
DB: 2 Gaps: 0

US-10-728-051-2 (1-171) x Q6PSU1 (1-175)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCTGCCCGCCAGCATCTGCGAGGCGAG 61
Db 19 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 38
QY 62 CAGTGGGAATCCAAAGGAGACAGAATGCCAGAGCGAGCTCGAGAGGGCGAACCTGAGG 121
Db 39 GlnTrpGluPheGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 58
QY 122 CCTGCGAGCAATCTCATGCGAGATCCAAAGTGCAGGATTCATATGAACGGGAC 181
Db 59 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 78
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCGGAGGCGCTGGA 241

```

```

Db 79 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 98
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACAGTTCAGAACCAACAAAGG 301
Db 99 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 118
QY 302 TCATGTGCGAGGCAATTCGAACAGATCATGAGAACCAAGAGCATAGTTGCGAGGAGG 361
Db 119 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 138
QY 362 CACACGAGCAACAGTTCAGAGGGAGCTCAGGAATTCCTGCTCAACAGTCCGGCCCTAGG 421
Db 139 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 158
QY 422 GCACCCAGCGTTCGCACTTGGACGTGCAAAAGTGGCGGAGAGACAGATAC 472
Db 159 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 175

RESULT 3
Q941R0
ID Q941R0 PRELIMINARY; PRT; 156 AA.
AC Q941R0;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Allergen II (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-21192158; PubMed=11295663; DOI=10.1067/mai.2001.113522;
RA Viquez O.M., Summer C.G., Dodo H.W.;
RT "Isolation and molecular characterization of the first genomic clone
of a major peanut allergen, Ara h 2."
RL J. Allergy Clin. Immunol. 107:713-717(2001).
DR EMBL; AY007229; AAK96887.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON TER 156
SQ SEQUENCE 156 AA; 18167 MW; 664D6832F91F36D1 CRC64;

Alignment Scores:
Pred. No.: 2.68e-63 Length: 156
Score: 819.00 Matches: 153
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.74% Indels: 0
DB: 2 Gaps: 0

US-10-728-051-2 (1-171) x Q941R0 (1-156)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCTGCCCGCCAGCATCTGCGAGGCGAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAATCCAAAGGAGACAGAATGCCAGAGCGAGCTCGAGAGGGCGAACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCTGCGAGCAATCTCATGCGAGATCCAAAGTGCAGGATTCATATGAACGGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCGGAGGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 83

```

```

QY 242 TCCTCTCAGACCAAGAGAGGCTTTCGATAGCTGAAACGAGTTTGAGAACACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 103
QY 302 TGCATGTGCCAGGCACTTGCACAGATCATGCAGAACCCAGAGCGATAGGTTGCAGGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTCTCAACAGATGCGGCTTAGG 421
Db 124 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACACAGGCTTGCAGCTTGGAGCTGCGAAAGTGGCGGC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156

RESULT 4
ID Q8GV20 PRELIMINARY; PRT; 172 AA.
AC Q8GV20;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Allergen Ara h 2.02.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Chatel J.-M., Bernard H., Orson F.M.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL: AY158467; AAN7576.1; -.
DR InterPro: IPR003612; AAI.
DR Pfam: PF00234; Tryp_alpha_aml1; 1.
DR SMART: SM00499; AAI; 1.
SQ SEQUENCE 172 AA; 20114 MW; B8BB91C8D8C143AB CRC64;

Alignment Scores:
Pred. No.: 6.04e-63 Length: 172
Score: 815.00 Matches: 155
Percent Similarity: 92.31% Conservative: 1
Best Local Similarity: 91.72% Mismatches: 1
Query Match: 64.43% Indels: 12
DB: 2 Gaps: 1

US-10-728-051-2 (1-717) x Q8GV20 (1-172)
QY 2 CTACCACTACTAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCGATCTGCGAGGAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAATCCCAAGGAGACAGAAAGATGCCAGAGCGCTCGAGAGGGCGAACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCCTGCGAGCAGCATCTCATGCGAGAGATCCAACTGACGAGGATTCATATGACGGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrglyArgAsp 63
QY 182 CCGTACAGCCCTAGTCAG-----GATCCG 205
Db 64 ProTySerProSerGlnAspProTySerProSerGlnAspProAspArgAspPro 83
QY 206 TACAGCCCTAGTCATATGATCGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGT 265
Db 84 TySerProSerProTySerAspArgGlyAlaGlySerSerGlnHisGlnGluArgCys 103
QY 266 TGCATAGCTGAACGAGTTTGAGAACAAACAAAGGTGTCATGTCGAGGCGCATTCGCAACAG 325
Db 104 CysAsnGluLeuAsnGluPheGluAsnGlnArgCysMetCysGluAlaLeuGlnGln 123

```

```

QY 326 ATCATGGAGAACCAAGAGCGATAGGTTGCAGGGGAGGCAACAGAGCAACAGTTCAAGAGG 385
Db 124 IleMetGluAsnGlnSerAspArgLeuGlnGlyArgGlnGlnGlnGlnPheLysArg 143
QY 386 GAGCTCAGGAACCTTGCTCAACAGTCGCGCTTAGGGCCACCAGCGTTGCGAGCTTGGAC 445
Db 144 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 163
QY 446 GTCGAAAGTGGCGGCGACAGACATAC 472
Db 164 ValGluSerGlyGlyArgAspArgTyr 172

RESULT 5
ID Q6PSU2 PRELIMINARY; PRT; 187 AA.
AC Q6PSU2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Seed storage protein SSPI (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Liao B., Li H., Lin X., Huang S.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL: AY581853; AAR00598.1; -.
DR InterPro: IPR003612; AAI.
DR Pfam: PF00234; Tryp_alpha_aml1; 1.
DR SMART: SM00499; AAI; 1.
DR NON_TER 1.
SQ SEQUENCE 187 AA; 21786 MW; 850AF22C94983FB8 CRC64;

Alignment Scores:
Pred. No.: 6.07e-63 Length: 187
Score: 815.00 Matches: 155
Percent Similarity: 92.31% Conservative: 1
Best Local Similarity: 91.72% Mismatches: 1
Query Match: 64.43% Indels: 12
DB: 2 Gaps: 1

US-10-728-051-2 (1-717) x Q6PSU2 (1-187)
QY 2 CTACCACTACTAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCGATCTGCGAGGAG 61
Db 19 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 38
QY 62 CAGTGGGAATCCCAAGGAGACAGAAAGATGCCAGAGCGCTCGAGAGGGCGAACCTGAGG 121
Db 39 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 58
QY 122 CCCTGCGAGCAACATCTCATGCGAGAGATCCAACTGACGAGGATTCATATGAAACGGGAC 181
Db 59 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrglyArgAsp 78
QY 182 CCGTACAGCCCTAGTCAG-----GATCCG 205
Db 79 ProTySerProSerGlnAspProTySerProSerGlnAspProAspArgAspPro 98
QY 206 TACAGCCCTAGTCATATGATCGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGT 265
Db 99 TySerProSerProTySerAspArgGlyAlaGlySerSerGlnHisGlnGluArgCys 118
QY 266 TGCATAGCTGAACGAGTTTGAGAACAAACAAAGGTGTCATGTCGAGGCGCATTCGCAACAG 325
Db 119 CysAsnGluLeuAsnGluPheGluAsnGlnArgCysMetCysGluAlaLeuGlnGln 138
QY 326 ATCATGGAGAACCAAGAGCGCATAGTTGTCAGGGGAGGCAACAGGAGCAACAGTTCAAGAGG 385

```





Db 131 ArgAlaProGlnArgCysAspLeuAspVal---SerGlyGlyArg 144

RESULT 9

Q8W251 PRELIMINARY; PRT; 144 AA.

ID Q8W251 AC Q8W251; 20, Created

DT 01-WAR-2002 (TRENBLrel. 20, Last sequence update)

DT 01-WAR-2002 (TRENBLrel. 20, Last sequence update)

DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)

DE Conglutinin (Fragment).

OS Arachis hypogaea (Peanut).

OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;

OC Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, rosids;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;

OC Arachis.

ON NCBI\_TaxID=3818;

OX [1]

RN R1

RP SEQUENCE FROM N.A.

RX PubMed=12582692;

RA Paik-Ro O.G., Seib J.C., Smith R.L.;

RL "Seed-specific, developmentally regulated genes of peanut.";

RL Theor. Appl. Genet. 104:236-240(2002).

DR EMBL; AF366561; AAL37561.1; -;

DR Pfam; PF00234; Tryp\_alpha\_aml1; 1.

DR SMART; SM00499; AAI; 1.

FT NON\_TER 1

SQ SEQUENCE 144 AA; 16790 MW; B776486138A852C3 CRC64;

Alignment Scores:

Pred. No.:	2,53e-30	Length:	144
Score:	443.50	Matches:	90
Percent Similarity:	69.68%	Conservative:	18
Best Local Similarity:	58.06%	Mismatches:	30
Query Match:	35.06%	Indels:	17
DB:	2	Gaps:	4

US-10-728-051-2 (1-717) x Q8W251 (1-144)

QY 5 ACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCATCTGCG---AGGCAG 61

Db 4 ThrilleLeuValAlaLeuLeuAlaLeuValLeuAlaHisAlaSerAlaMetArgArg 23

QY 62 CAGTGGGAACCTCCAGAGAGAC---AGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTG 118

Db 24 GluArgGlyArgGlnGlyAspSerSerCysGluArgGlnValAspArgValAsnLeu 43

QY 119 AGGCCCTCGAGCAACATCTCATGCAGAGATCAACGTCGACGAGGATTCATATGAACGG 178

Db 44 LysProCysGluGlnHisIleMetGlnArgIleMetGlyGlnGlnIlyrAsp--- 62

QY 179 GACCCGTTACAGCCCTAGTTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCT 238

Db 63 -----SerTyrAspIleArgSerThr 69

QY 239 GGATCCTCTACGACCAGAGAGGTGTTCATAGCTGAACGAGTTTGAGAACACCAA 298

Db 70 ArgSerSerAspGlnGlnArgCysCysAspGluLeuAspGlnMetGluAsnThrGlu 89

QY 299 AGGTGCATGTCGAGGCATTCACACATCATCGAGAACCCAGAGCCATAGTTGTCAGGGG 358

Db 90 ArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnCysAspArgLeuGlnAsp 109

QY 359 AGGCCAACAGGAGCAACAGTTCACAGAGGAGCTCAGGAACCTTGCTCAACAGTCGCGCCTT 418

Db 110 ArgGlnMetValGlnGlnPheLysArgGluLeuMetAsnLeuProGlnGlnCysAsnPhe 129

QY 419 AGGGCACCCAGCGTTGCGACTTCGGACGTCGAAAGTCGGCGGCGAGA 463

Db 130 ArgAlaProGlnArgCysAspLeuAspVal---SerGlyGlyArg 143

RESULT 10

Q9SQG5 PRELIMINARY; PRT; 129 AA.

ID Q9SQG5

AC Q9SQG5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Allergen Arah6 (Fragment).  
 GN Name=Ara h 6;  
 OS Arachis hypogaea (Peanut).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
 OC Arachis.  
 OX NCBI\_TaxID=3818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seed;  
 RX MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;  
 RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.M.;  
 RT "Selective cloning of peanut allergens, including profilin and 2S  
 albumins, by phase display technology.";  
 RL Int. Arch. Allergy Immunol. 119:265-274 (1999).  
 DR EMBL; AF092846; AAD56337.1; -.  
 DR InterPro; IPR003612; AAI.  
 DR Pfam; PF00234; Tryp\_alpha\_aml; 1.  
 DR SMART; SM00499; AAI; 1.  
 FT NON TER  
 SQ SEQUENCE 129 AA; 14929 MW; 4048D3418BB9EE40 CRC64;  
 Alignment Scores:  
 Pred. No.: 1,7e-25 Length: 129  
 Score: 388.50 Matches: 80  
 Percent Similarity: 66.43% Conservative: 15  
 Best Local Similarity: 55.94% Mismatches: 31  
 Query Match: 30.71% Indels: 17  
 DB: 2 Gaps: 4  
 US-10-728-051-2 (1-717) x Q9SQG5 (1-129)  
 QY 41 GCCACGATCTGGG---AGCAGCAGTGGGAACCTCCAGAGGAC---AGAAGATGCCAG 94  
 DB 1 AlahisalaSerAlaMetArgGluArgGlyArgGlnGlyAspSerSerCysGlu 20  
 QY 95 AGCCAGCTCGAGAGGGCGAACCTGAGGCCCTGCGAGCAACATCTCATGCAAGATCCAA 154  
 DB 21 ArgGlnValAspGlyValAlaSerLeuLysProCysGluGlnHisLeuMetGlnArgMet 40  
 QY 155 CGTGACGAGGATTCATATGAACGGGACCCGTCACAGCCCTAGTCAGGATCGTACAGCCCT 214  
 DB 41 GlyGluGlnGluGlnTyrAsp----- 47  
 QY 215 AGTCCATATCGGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGTTCATGAG 274  
 DB 48 ---SerTyrAsnPheGlySerThrArgSerSerAspGlnGlnArgCysAspGlu 66  
 QY 275 CTGAACAGTTCGAGAACCAACAAAGTGTGATGTGCGAGGCATTGCAACAGATCATGGAG 334  
 DB 67 LeuAsnGluMetGluAsnThrGlnArgCysMetCysGluAlaLeuGlnGlnMetGlu 86  
 QY 335 AACAGAGCATAGTGTTCAGGGAGGCGCAACAGAGCAACAGTTCAGAGGAGGTCCAGG 394  
 DB 87 AsnGlnCysAspGlyLeuGlnAspArgGlnMetValGlnHisPheLysArgGluLeuMet 106  
 QY 395 AACTTCCTCAACAGTGGCGCTTAGGGCAACCAAGCTGTGGACCTGGACGTGCAAGT 454  
 DB 107 AsnLeuProGlnGlnCysAsnPheGlyAlaProGlnArgCysAspLeuAspVal---Ser 125  
 QY 455 GGCGGCAGA 463  
 DB 126 GlyGlyArg 128  
 RESULT 11  
 Q9SQH1  
 ID Q9SQH1 PRELIMINARY; PRT; 160 AA.  
 AC Q9SQH1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Allergen Arah6 (Fragment).  
 GN Name=Ara h 6;  
 OS Arachis hypogaea (Peanut).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
 OC Arachis.  
 OX NCBI\_TaxID=3818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seed;  
 RX MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;  
 RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.M.;  
 RT "Selective cloning of peanut allergens, including profilin and 2S  
 albumins, by phase display technology.";  
 RL Int. Arch. Allergy Immunol. 119:265-274 (1999).  
 DR EMBL; AF092846; AAD56337.1; -.  
 DR InterPro; IPR003612; AAI.  
 DR Pfam; PF00234; Tryp\_alpha\_aml; 1.  
 DR SMART; SM00499; AAI; 1.  
 FT NON TER  
 SQ SEQUENCE 129 AA; 14929 MW; 4048D3418BB9EE40 CRC64;  
 Alignment Scores:  
 Pred. No.: 1,7e-25 Length: 129  
 Score: 388.50 Matches: 80  
 Percent Similarity: 66.43% Conservative: 15  
 Best Local Similarity: 55.94% Mismatches: 31  
 Query Match: 30.71% Indels: 17  
 DB: 2 Gaps: 4  
 US-10-728-051-2 (1-717) x Q9SQH1 (1-129)  
 QY 41 GCCACGATCTGGG---AGCAGCAGTGGGAACCTCCAGAGGAC---AGAAGATGCCAG 94  
 DB 1 AlahisalaSerAlaMetArgGluArgGlyArgGlnGlyAspSerSerCysGlu 20  
 QY 95 AGCCAGCTCGAGAGGGCGAACCTGAGGCCCTGCGAGCAACATCTCATGCAAGATCCAA 154  
 DB 21 ArgGlnValAspGlyValAlaSerLeuLysProCysGluGlnHisLeuMetGlnArgMet 40  
 QY 155 CGTGACGAGGATTCATATGAACGGGACCCGTCACAGCCCTAGTCAGGATCGTACAGCCCT 214  
 DB 41 GlyGluGlnGluGlnTyrAsp----- 47  
 QY 215 AGTCCATATCGGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGTTCATGAG 274  
 DB 48 ---SerTyrAsnPheGlySerThrArgSerSerAspGlnGlnArgCysAspGlu 66  
 QY 275 CTGAACAGTTCGAGAACCAACAAAGTGTGATGTGCGAGGCATTGCAACAGATCATGGAG 334  
 DB 67 LeuAsnGluMetGluAsnThrGlnArgCysMetCysGluAlaLeuGlnGlnMetGlu 86  
 QY 335 AACAGAGCATAGTGTTCAGGGAGGCGCAACAGAGCAACAGTTCAGAGGAGGTCCAGG 394  
 DB 87 AsnGlnCysAspGlyLeuGlnAspArgGlnMetValGlnHisPheLysArgGluLeuMet 106  
 QY 395 AACTTCCTCAACAGTGGCGCTTAGGGCAACCAAGCTGTGGACCTGGACGTGCAAGT 454  
 DB 107 AsnLeuProGlnGlnCysAsnPheGlyAlaProGlnArgCysAspLeuAspVal---Ser 125  
 QY 455 GGCGGCAGA 463  
 DB 126 GlyGlyArg 128  
 RESULT 11  
 Q9SQH1  
 ID Q9SQH1 PRELIMINARY; PRT; 160 AA.  
 AC Q9SQH1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Allergen Arah6 (Fragment).  
 GN Name=Ara h 7;  
 OS Arachis hypogaea (Peanut).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
 OC Arachis.  
 OX NCBI\_TaxID=3818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Virginia; TISSUE=Seed;  
 RX MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;  
 RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.M.;  
 RT "Selective cloning of peanut allergens, including profilin and 2S  
 albumins, by phase display technology.";  
 RL Int. Arch. Allergy Immunol. 119:265-274 (1999).  
 DR EMBL; AF091737; RAD56719.1; -.  
 DR InterPro; IPR003612; AAI.  
 DR SMART; SM00499; AAI; 1.  
 SQ SEQUENCE 160 AA; 18417 MW; 9F9E4CEE68808D4C CRC64;  
 Alignment Scores:  
 Pred. No.: 1,02e-17 Length: 160  
 Score: 300.00 Matches: 72  
 Percent Similarity: 61.82% Conservative: 30  
 Best Local Similarity: 43.64% Mismatches: 33  
 Query Match: 23.72% Indels: 31  
 DB: 2 Gaps: 7  
 US-10-728-051-2 (1-717) x Q9SQH1 (1-160)  
 QY 2 CTCACCATACTAGTAGCCCTC-----GCCCTTTTCCTCTCGTGCC----- 43  
 DB 5 LeuSerIleLeuValAlaLeuLeuGlyAlaLeuValAlaSerAlaThrArgTrp 24  
 QY 44 -----CAGCATCTCGAGGCGAGCAGTGGGAACTC-----CAAGGACAGAGA 88  
 DB 25 AspProAspArgGlySerArgGlySerArgTrpAspAlaProSerArgGlyAspGln 44  
 QY 89 TGCCAGAGCCAGCTCGAGAGGCGAACTGAGGCCCTCGAGCAACATCTCATGCGAGAAG 148  
 DB 45 CysGlnArgGlnLeuGlnArgAlaAsnLeuArgProCysGluGluHisMetArgArg 64  
 QY 149 ATCCAACTGACGAGGATTCATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTAC 208  
 DB 65 ValGluGlnGluGlnGluGln-----GluGlnAspGluTyr----- 76  
 QY 209 AGCCCTAGTCCATATGATCGGAGAGGCGCTGATCCTCTCAG----- 250  
 DB 77 -----ProTyrSerArgArgGlySerArgGlyArgGlnProGlyGluSerAspGlu 93  
 QY 251 ---CACCAGAGAGGTGTTCATGAGCTGAGTGAACGAGTTTGAGAACCAAGAGTGTGATG 307  
 DB 94 AsnGlnGluGlnArgCysCysAsnGluLeuAsnArgPheGlnAsnAsnGlnArgCysMet 113  
 QY 308 TGCCAGGCTATTGCAACAGATCATGAGAACCCAGAGCCGATAG-----GTTGCGAGGAGGCA 363  
 DB 114 CysGlnAlaLeuGlnGlnIleLeuGlnAsnGlnSerPheTrpValProAlaGlyGlnGlu 133  
 QY 364 ACAGGAGCAACAGTTCAGAGGAGGCTCAGGAAGTTCGCTCAACAGTTCGCGCCCTAGGCG 423  
 DB 134 ProValAlaSerAspGlyGlyGlyAlaGlnGluLeuAlaProGluLeuArgValGlnVal 153  
 QY 424 ACCACAGCGTTCGCA 438  
 DB 154 ThrLysProLeuArg 158  
 RESULT 12  
 Q647G8  
 ID Q647G8 PRELIMINARY; PRT; 158 AA.

AC Q647G8; 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE 2S protein 2.  
 OS Arachis hypogaea (Peanut).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
 OC Arachis.  
 OX NCBI\_TaxID=3818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yan Y., Wang L., Huang S.;  
 RT "CDNA clone of peanut seed storage protein gene.";  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY72691; AAU21496.1; -.  
 SQ SEQUENCE 158 AA; 18455 MW; 32AD4D4061544948 CRC64;

Alignment Scores:  
 Pred. No.: 1.25e-17 Length: 158  
 Score: 299.00 Matches: 75  
 Percent Similarity: 63.12% Conservative: 26  
 Best Local Similarity: 46.88% Mismatches: 39  
 Query Match: 23.64% Indels: 21  
 DB: 2 Gaps: 7

US-10-728-051-2 (1-717) x Q647G8 (1-158)

QY 2 CTCACCATCTAGTACGCCCTC-----GCCCTTTCTCTCTCTCGTCC----- 43  
 Db 4 LeuSerIleuValAlaLeuLeuGlyAlaLeuValAlaSerAlaThrArgTyr 23  
 QY 44 -----CACGCATCTCGGAGCGACGACGAGTGGGAACCTC-----CAAGGAGACAGAGA 88  
 Db 24 AspProAspArgGlySerArgGlyLeuArgTyrAspAlaProSerArgGlyAspAspGln 43  
 QY 89 TGCAGAGCCAGCTCGAGAGGCGAACCTGAGGCCCTCGAGCAACATCTCATGCGAGAAG 148  
 Db 44 CysGlnArgGlnLeuGlnArgAlaAsnLeuArgProCysGluGluHisIleArgGlnArg 63  
 QY 149 ATCCACGTCAGCAGGATTTCATATGACGGGACCCGTACAGCCCTAGTCAGGATCCGTAC 208  
 Db 64 ValGluGlnGluGlnGluGln-----GluGlnAspGluTyr-----ProTyr 77  
 QY 209 AGCCCT-----AGTCCATATGATCGGAGAGCGCTGATCTCTCAGCACCAAGAGAGG 262  
 Db 78 SerGlnArgGlySerArgGlyArgArgProGlyGluSerAspGluAspGlnGlnArg 97  
 QY 263 TGTTCATAGCTGAACGAGTTTGAGAACCAACAAAGGTGCGATGTCGAGGCGATTGCAA 322  
 Db 98 CysCysAsnGluLeuAsnArgPheGlnAsnAsnGlnArgCysMetCysGlnAlaLeuGln 117  
 QY 323 CAGATCATGAGAACCAAGCGATAGTTGTTG-----CAGGGAGGCAACAGGACCAAGTT- 378  
 Db 118 GlnIleuLeuAsnGlnSerPheArgPheGlnGlnAspArgSerGlnLeuHisGlnAsn 137  
 QY 379 CAAGAGGAGCTCAGGAACCTGCTCAACAGTGGCGCTTTAGGGCACACAGCGCTTGGCA 438  
 Db 138 GlyGluGlyAlaGlnGluLeuAlaProGluLeuArgValGlnValThrLysProLeuArg 157

RESULT 13  
 Q99235  
 ID Q99235 PRELIMINARY; PRT; 153 AA.  
 AC Q99235;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Conglutin delta.  
 OS Lupinus angustifolius (Narrow-leaved blue lupine).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.

OX NCBI\_TaxID=3871;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=unicrop; TISSUE=Cotyledon;  
 RX MEDLINE=91355912; PubMed=2103479;  
 RA Gayler K.R., Koliavas S., Macfarlane A.J., Lilley G.G., Baldi M.,  
 RA Biagrover R.J., Johnson E.D.;  
 RT "Biosynthesis, cDNA and amino acid sequences of a precursor of  
 RT conglutin delta, a sulphur-rich protein from *Lupinus angustifolius*.";  
 RL Plant Mol. Biol. 15:879-893(1990).  
 DR EMBL; X53523; CAA37598.1; -.  
 DR PIR; S12404; A33090.  
 DR InterPro; IPR003612; AAI.  
 DR Pfam; PF00234; Tryp\_alpha\_amy1; 1.  
 DR SMART; SM00499; AAI; 1.  
 SQ SEQUENCE 153 AA; 17785 MW; A45B7025353F0769 CRC64;

Alignment Scores:  
 Pred. No.: 8.7e-16 Length: 153  
 Score: 278.00 Matches: 62  
 Percent Similarity: 55.00% Conservative: 26  
 Best Local Similarity: 38.75% Mismatches: 46  
 Query Match: 21.98% Indels: 26  
 DB: 2 Gaps: 5

US-10-728-051-2 (1-717) x Q99235 (1-153)

QY 2 CTCACCATCTAGTACGCCCTC-----GCCCTTTCTCTCTCTCGTCCCGCAGG 58  
 Db 4 LeuThrIleuLeuAlaLeuValAlaLeuValAlaValHisThrSerAlaPhe 23  
 QY 59 CAGCAGTGGGAACCTCCAGGAGACAGAGATGCCAGGCGAGCTCGAGAGGGCAACCTG 118  
 Db 24 Gln-----SerSerLysGlnSerCysLysArgGlnLeuGlnValAsnLeu 39  
 QY 119 AGCCCTCGGAGCAACATCTCATGCGAGAGATCCACGTCACGAGGATTCATATGACGG 178  
 Db 40 ArgHisCysGluAsnHisIleAlaGlnArgIleGlnGlnGlnGluGluAsp 59  
 QY 179 GACCCGTACAGC-----CTAGCTCAGGATCCG 205  
 Db 60 HisAlaLeuLysLeuArgGlyIleLysHisValIleuArgHisArgSerGlnGlu 79  
 QY 206 TACAGCCCTAGTCCATATGATCGGAGAGCGCTGGATCTCTCAGCACCAAGAGAGTGT 265  
 Db 80 TyrSerGlu-----GluSerGluGluLeuAspGlnCys 90  
 QY 266 TGCATGAGCTGACAGTGTGAGAACCAACCAAGGTGATGTCGAGGCGATTGCCACAG 325  
 Db 91 CysGluGlnLeuAsnGluLeu-----AsnSerGlnArgCysGlnCysArgAlaLeuGlnGln 109  
 QY 326 ATCATCGAGAACCAAGCGATAGTTGCGAGGGAGGCAACAGGAGCAACAGTTCAAGAGG 385  
 Db 110 IleTyrGluSerGlnSerGlnGlnCysGluGlySerGlnGlnGlnGlnGluGln 129  
 QY 386 GAGCTCAGGAACCTGCTCAACAGTCCGCGCTTAGGGCAACACAGCGTTGCGATTGGAC 445  
 Db 130 GluLeuGluLysLeuProArgThrCysGlyPheGlyProLeuArgCysAspValAsn 149

RESULT 14  
 2SS SOYBN  
 ID 2SS SOYBN STANDARD; PRT; 158 AA.  
 AC P19594;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE 2S albumin precursor (GM2S-1) [contains: 2S albumin small chain  
 DE (Aspartic acid-rich peptide); 2S albumin large chain (8 kDa  
 DE methionine-rich protein) (8 kDa MRP)].  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

```

OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hodgson 78; TISSUE=Cotyledon;
RA Galvez A.F., Revilla M.J.R., de Lumen B.O.;
RT "A novel methionine-rich protein from soybean cotyledon: cloning and
RL characterization of cDNA.";
RL (er) Plant Gene Register PGR97-103.
RN [2]
RP SEQUENCE OF 22-64.
RP TISSUE=Seed;
RX MEDLINE=87280104; PubMed=3611081;
RA Odani S., Koide T., Ono T.;
RT "Amino acid sequence of a soybean (Glycine max) seed polypeptide
RL having a poly(L-aspartic acid) structure.";
RN [3]
RP SEQUENCE OF 82-96.
RA Revilla M.J., Galvez A.F., Krenz D.C., de Lumen B.O.;
RT "An 8 kDa methionine-rich protein (MRP) from soybean (Glycine max)
RL cotyledon: identification, purification and N-terminal sequence.";
RN [4]
RP J. Agric. Food Chem. 44:2930-2935(1996).
CC -!- FUNCTION: This is a 2S seed storage protein.
CC -!- SUBUNIT: The protein consists of two chains linked by 2 disulfide
CC bonds.
CC -!- SIMILARITY: Belongs to the 2S seed storage albumins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF005030; AAB71140.1; -.
DR PIR; T05710; T05710; AAI.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_aml1; 1.
DR SMART; SM00499; AAI; 1.
DR Direct protein sequencing; Seed storage protein; Signal.
KW SIGNAL
FT CHAIN 1 21
FT CHAIN 22 64 2S albumin small chain.
FT PROPEP 65 81
FT CHAIN 82 158 2S albumin large chain.
FT SITE 54 56 Cell attachment site (Potential).
FT DOMAIN 56 64 Poly-Asp.
FT DOMAIN 88 91 Poly-Glu.
SQ SEQUENCE 158 AA; 18460 MW; C1F42723B0F62D81 CRC64;

Alignment Scores:
Pred. No.: 1.03e-12 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.49% Conservative: 31
Best Local Similarity: 36.36% Mismatches: 55
Query Match: 19.21% Indels: 12
DB: 1 Gaps: 6

US-10-728-051-2 (1-717) x 2SS_SOYBN (1-158)
QY 5 ACCATAGTAGTCCCTCGCCCTTCTCTCTCGCTGCCAGCATCTCGAGGACGAC 64
Db 5 ThrLeuLeuLeuSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGAACTCAAGAGACAGAGAGATCCAGACGACCTCGAGAGGCGGCACTGAGGCC 124
Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42
QY 125 TGGAGCAACATCTTCATGCAGAGATCCAA-----CGTGACGAGATTTCATGACGG 178
Db 43 CysGlnHisLeuMetGlnLysIleGlnGlyArgGlyAspAspAspAsp 62
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCT 238

```

```

Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgAsnGlu 82
QY 239 GGATCTCTCAG-----CACCAGAGAGGTGTTCGAATGAGCTGAAC 280
Db 83 GlyLysAspGluAspGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102
QY 281 GACTTTGAGACAACCAAGAGTTCATGTCGAGGCAATTCACAGATCATGAGAACCAAG 340
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 AGCATATAGTTTCAGCGGAGGCAACAGAGCAACAGTTTCAAGAGGAGCTCAGGACTTG 400
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysLysMetGluLysGluLeuLeuLeu 140
QY 401 CCTCAACAGTGGCGCTTAGGGACCAACACAGCGTTCGACTTG 442
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 15
Q9ZNZ4
ID Q9ZNZ4 PRELIMINARY; PRT; 155 AA.
AC Q9ZNZ4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Napin-type 2S albumin 1 precursor.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=variety PIONEER 9341; TISSUE=Seed;
RA Jung R., Hastings C., Coughlan S.J., Hu W.-N.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U71194; AAD09630.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_aml1; 1.
DR SMART; SM00499; AAI; 1.
KW Signal.
FT SIGNAL 1 20 Potential.
FT CHAIN 21 155 napin-type 2S albumin 1.
SQ SEQUENCE 155 AA; 17835 MW; 2C15E55E936301E3 CRC64;

Alignment Scores:
Pred. No.: 1.4e-12 Length: 155
Score: 241.50 Matches: 58
Percent Similarity: 57.14% Conservative: 30
Best Local Similarity: 37.66% Mismatches: 53
Query Match: 19.09% Indels: 13
DB: 2 Gaps: 7

US-10-728-051-2 (1-717) x Q9ZNZ4 (1-155)
QY 2 CTCACATAGTAGTCCCTCGCCCTTCTCTCTCGCTGCCAGCATCTCGAGGACG 61
Db 4 LeuThrIleLeuLeuAlaLeuLeuPheIle-----AlaHisThrCysCysAlaSer 21
QY 62 CAGTGGAACTCCAGGAGAGAGATCCAGAGCCAGCTCGAGAGGCGCAACTGAGG 121
Db 22 LysTrpGlnGlnHisGlnGlnGlnSerCysArgGluGlnLeuLysGlyIleAsnLeuAsn 41
QY 122 CCTCGAGCAACATCTTCATGCAGAGATCCAA-----CGTGACGAGATTTCAT 172
Db 42 ProCysGlu---HisIleMetGlnLysIleGlnAlaGlyArgArgGlyGluAspGlySer 60
QY 173 GAACGGGACCGGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATGATGCGGAGA 232
Db 61 AspGluAspHisIleLeuLeuLeuArgThrMetPro---GlyArgIleAsnTyrIleArgLys 79
QY 233 GCGCTGGATCCTCTCAG-----CACCAGAGAGGTGTTCGAATGAGCTGNAC 280

```

```

Db      80 LysGluGlyLysGluGluGluGluGluGlyHisMetGlnLysCysCysSerGluMetSer 99
QY      281 GAGTTTGAGAAACCAAGTGCATGTGCGAGGCATTGCAACAGATCATGTGAGAAACCCAG 340
Db      100 GluLeuLys---SerProIleCysGlnCysLysAlaLeuGlnLysIleMetAspAsnGln 118
QY      341 AGCGATAGTTGCGAGGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTG 400
Db      119 SerGluGlnLeuGluGlyLys--GluLysLysGlnMetGluArgGluLeuMetAsnLeu 137
QY      401 CCTCAACAGTGGCGCCTTAGGGCACACACGCTTGGGACTTG 442
Db      138 AlaIleArgCysArgLeuGlyProMetIleGlyCysAspLeu 151

```

Search completed: April 8, 2005, 05:25:11 ~  
Job time : 128.066 secs

**This Page Blank (uspto)**

Result No.	Score	Query Match	Length	DB	ID	Description
1	1502	55.8	481	2	S04605	glycinin G3 - soyb
2	1499.5	55.7	485	2	S11002	glycinin G2 precu
3	1496	55.6	495	2	S10851	glycinin G1 precu
4	1492.5	55.5	485	1	FWSYG1	glycinin chain A2B
5	1466	54.5	495	1	FWSYG2	glycinin chain A1a
6	1457.5	54.2	498	2	S44294	legumin A precursor
7	1455	54.1	520	2	S08237	legumin A2 precursor
8	1440	53.5	517	1	FWPMLA	legumin A precursor
9	1431.5	53.2	482	2	S49877	legumin A precursor
10	1428	53.1	484	2	S11003	glycinin G3 precu
11	1414.5	52.6	500	2	S14393	legumin A2 precursor
12	1384	51.4	507	2	T06452	probable legumin A
13	1375	51.1	497	2	S14392	legumin A1 precursor
14	1073.5	39.9	551	2	S51941	prunin 1 precursor



```

QY 121 GAATGCGCGCGCTGCGCTCTCTCGCTTAGTCTCCGCGCAACGCGCTTCGTAGCGCT 180
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
QY 181 TTCTACTCCAAATGCTCCAGGAGATCTTCATCAGCAAGAGAGGATCTTTGGTTG 240
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
QY 241 ATATTCTCTGTTGCTCTAGACTACTATGAAGAGCTTCACACAAAGTCTCGATCTCAG 300
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnGlyGly-----Gln 119
QY 301 TCCCAAGACCAACAAGACGCTCCCAAGGAGAAAGCAACCAAGCAACGACGAGATAGT 360
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 SerSerArgPro-----GlnAspArg 126
QY 361 CACCAAGAGTGCACCGTTTCGATGAGGTGATCTATTTCAGTTCCACCGTGTGTCT 420
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
127 HisGlnLysIleTyrHisPheArgGluGluIleAlaValProThrGlyPheAla 146
QY 421 TTCTGGCTCTACACCAACCAACGACACTGATGTTGTTGCTGTTCTTCTACTGACACCAAC 480
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
147 TyrTyrMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
QY 481 AACACGACACACGCTTGATAGTCCCGCAGAGATTCATTTGCTGGGAACACGGAG 540
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY 541 CAAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAGAGCAGAGAGAGCTTACCATAT 600
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly-----199
QY 601 AGCCCATACAGCCCGCAAGTACGCTAGACAAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
200 -----Gly 200
QY 661 CACCAAGCGCAGAGAACGAGCAGGACCAAGAAAGAAAGAAAGAGTGGAAACATCTTC 720
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGluGluAsnGlnGlyGlySerLeu 220
QY 721 AGCGGCTTCACGCGGAGTCTCCGGAACAAAGCCTCCAGGTTGACACAGACAGATAGT 780
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
221 SerGlyPheAlaProGluPheLeuGluHisAlaPheVal---AspArgGlnIleVal 239
QY 781 CAAACCTAAGAGCGGAGCCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGGGGA 840
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
240 ArgLysLeuGlnGlyGluAsnGluGluGluLysGlyAlaIleValThrValLysGly 259
QY 841 GGCCTCAGATCTTGAGCCCA-----GATAGAAAGAGAGCTGCCAGCAGAGAGAG 891
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
260 GlyLeuSerValIleSerProThrGluGluGlnGlnGlnArgProGluGluGlu 279
QY 892 GAATACGATGAAGATGAATATGATACGATGAAGAGATAGAGGCGTGGCAGGGAAGC 951
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
QY 952 AGAGCAGGCGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAGAACATT 1011
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
QY 1012 GTTAGAAACAGATCCCTCAGCATCTACACCTCAGCTGGTTCATCTCAAACTGCCAAC 1071
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
QY 1072 GATCTCAACCTCTAATACTAGTGGCTTGGACCTAGTGGTGAATATGAAATCTCTAC 1131
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
QY 1132 AGAATGCAATGTTTTCGCTACTACACCAACGACGACAGCATCATATATCATGATG 1191
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
QY 1192 AGGGGACGGGCTCAGCTGCAAGTCGCGGACAGCAACGCGCAACAGAGTGTACACGAGGAG 1251

```

```

Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1252 CTTCAAGAGGTCACGCTGCTTGTGGTCCACAGAACTTCGCGCTCGCTGGAAAGTCCAG 1311
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
QY 1312 ACGGAACTTCGAATACGTCGTCATTCAGACAGACTCAAGGCCAGCATACCAACCTC 1371
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
QY 1372 GCCGCTGAAAACCTCCGTCATAGATAACCTCCGCGAGGAGTGTGCAATTCATATGCG 1431
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
QY 1432 CTCCTCAAGGAGCAGCAGGAGGAGCTTAAAGACACAAACCTTCAGTTCCTCGTTCCA 1491
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
QY 1492 CCGTCTCAGCAGCTCCGAGGCGCTGTGGCT 1521
Db   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
472 ProLysGluSerGlnArgArgValValAla 481

```

## RESULT 2

S11002

glycinin G2 precursor - soybean

N:Alternate names: glycinin A2B1a

C:Species: Glycine max (soybean)

C&gt;Date: 21-Nov-1993 #sequence\_revision 19-Jan-1996 #text\_change 09-Jul-2004

C:Accession: S11002; S04604; A26990

R:Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallion, B.J.; Fischer, R.L.;

Plant Cell 1, 313-328, 1989

A:Title: Characterization of the glycinin gene family in soybean.

A:Reference number: S10851; MUID:92393391; PMID:2485233

A:Accession: S11002

A:Molecule type: DNA

A:Residues: 1-485 &lt;NIE&gt;

A:Cross-references: UNIPROT:P04405

A:Experimental source: variety Dare

R:Thanh, V.H.; Tumer, N.E.; Nielsen, N.C.

Nucleic Acids Res. 17, 4387, 1989

A:Title: The glycinin G2 gene from soybean.

A:Reference number: S04604; MUID:89296499; PMID:2740230

A:Accession: S04604

A:Molecule type: DNA

A:Residues: 1-485 &lt;THA&gt;

A:Cross-references: EMBL:X15122; NID:g18636; PIDN:CAA33216.1; PID:g18637

A:Experimental source: variety Dare

R:Fukazawa, C.; Momma, T.; Higuchi, W.; Uda, K.

Nucleic Acids Res. 15, 8117, 1987

A:Title: Complete nucleotide sequence of the gene encoding a glycinin A2B1a subunit prec

A:Reference number: A26990; MUID:88040439; PMID:3671077

A:Accession: A26990

A:Molecule type: DNA

A:Residues: 1-485 &lt;FUK&gt;

A:Cross-references: GB:X02806

C:Genetics:

A:Gene: Gv2

A:Introns: 93/1; 177/3; 356/3

C:Superfamily: glycinin

C:Keywords: storage protein

F:1-18/Domain: signal sequence #status predicted &lt;SIG&gt;

F:19-485/Product: glycinin G2 #status predicted &lt;MAT&gt;

## Alignment Scores:

Pred. No.:	1-578-103	Length:	485
Score:	1499.50	Matches:	297
Percent Similarity:	72.02%	Conservative:	71
Best Local Similarity:	58.12%	Mismatches:	94
Query Match:	55.72%	Indels:	49
DB:	2	Gaps:	6

US-10-728-051-3 (1-1524) x S11002 (1-485)



















QY	1	CGGCAGCAACCGGAGGAGAACCGCTGCAGTTCCAGCGCTCAATGCGCAGACACCTTCAC	240
DB	23	ArgGluGlnSerGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuGluProAsp	102
QY	61	AATCGCATTGAATCAGAGCGCGGTTCACATTGAGCTTGGAAACCCCAACAACACGAGAGTTC	120
DB	43	AsnArgileGluSerGluGlyGlyLeuileGluThrTrpAsnProAsnAsnArgGlnPhe	62
QY	121	GAATGCGCGGGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCCCAACGGCCCTTCGTAGGCCT	180
DB	63	ArgCysAlaGlyValAlaLeuSerArgValThrLeuGlnArgAsnAlaLeuArgArgPro	82
QY	181	TTCTACTCCAAATGTCCTCCAGGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGGTTG	240
DB	83	TyrTrpSerAsnAlaProGlnGluileTyrlleGlnGlnGlyAsnGlyTyrPheGlyVal	102
QY	241	ATATTTCCTGGTTGCTCTAGACACTATGAAAGCCCTCACACAAAGTCGTTCGATCTCAG	300
DB	103	ValPheProGlyCysProGluThrPheGluGluPro-----GlnGluSer	117
QY	301	TCCCAAGACCCACCAAGACGCTCTCCAAGGAGAGACCAAGCCAAACAGACAGATAGT	360
DB	118	GluGlnArgGluArgArgArg-----TyrArgAspSer	128
QY	361	CACCAAGAGGTGCACCGTTTCGATGAGGGTGATCTCAATTGCATTTCCACCGGTGTGCT	420
DB	129	HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyAsnVal	148
QY	421	TTTCGGCTCTACACGACCGACACATGATGTTGTGCTGTTTCTTCTACTGACCAAC	480

## RESULT 10

S11003  
glycinin G3 precursor - soybean  
C;Species: Glycine max (soybean)  
C;Date: 21-Nov-1993 #sequence\_revision 19-Jan-1996 #text\_change 03-May-1996  
C;Accession: S11003  
R;Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallan, B.J.; Fischer, R.L.;  
Plant Cell 1, 313-328, 1989  
A;Title: Characterization of the glycinin gene family in soybean.  
A;Reference number: S10851; MUID:92393391; PMID:2485233  
A;Accession: S11003  
A;Molecule type: DNA  
A;Residues: 1-484 <NIE>  
A;Experimental source: variety Dare  
C;Keywords: glycinin  
C;Superfamily: storage protein  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-484/Product: glycinin G3 #status predicted <MAT>  
  
Alignment Scores:  
Pred. No.: 3.25e-98 Length: 484  
Score: 1428.00 Matches: 290  
Percent Similarity: 70.33% Conservative: 68  
Best Local Similarity: 56.97% Mismatches: 95  
Query Match: 53.07% Indels: 56  
DB: 2 Gaps: 9  
  
US-10-728-051-3 (1-1524) x S11003 (1-484)  
  
QY 1 CGGCAGCAACCGGAGGAGAACGGCTCCAGTTCAGCGCTCAATGCGCAGACCTGAC 60  
DB 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42  
QY 61 AATCGCATTAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120  
DB 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62  
QY 121 GAATGCCCGCGCTCGCTCTCTCGTTAGTCTCCGCGCAACGGCTTCGTAGGCCT 180  
DB 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82  
QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGAAAGGGGATCTTTGGGTG 240  
DB 83 SerTyrThrAsnAlaProGlnGluIleTyrlleGlnGlnGlySerGlyIlePheGlyMet 102  
QY 241 ATATTCCCTGGTGTCTTAGACACTATGAGAGGCTCACACAGAGTCTGTCGATCTCAG 300  
DB 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119  
QY 301 TCCCAAGACACCAAGACGCTCTCCAGGAGAGAGACCAAGGCCAACAGCAACGAGATAGT 360  
DB 120 SerSerArgPro-----GlnAspArg 126  
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCACCGGTGTGCT 420  
DB 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146  
QY 421 TTCTGGCTCTAACAGCAGCAGACACTGATGTGTGTGCTGTTCTCTTACTGACCAAC 480  
DB 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166  
QY 481 AACCAACGACACCAAGCTTGATTCAGTTCGCCAGGAGATTCAATTTGGCTGGGAACGAG 540  
DB 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186  
QY 541 CAAGAGTTCTTAAGGTACCAGCAACAAAGCAGCAACAAAGCAGCAGCAAGAGCTTACCATAT 600  
DB 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199  
QY 601 AGCCCATACACCGCGCAAAAGTTCAGCCTTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGGA 660

Db 200 -----|||  
QY 661 CAGCAGACCGCAGACAGACGAGCAGCAAGAGAAAGAAACGAGGTGGAACATCTTC 720  
Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlySerIleLeu 220  
QY 721 AGCGGCTTACCGCGAGTCTCTGGAACAGCCTTCCAGTTTCCAGCAGACAGATAGTG 780  
Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239  
QY 781 CAAACCTTAAGAGCGCAGACCGAGAGTGAAGAGAGAGAGAGCATTCTGACAGTGAGGGA 840  
Db 240 ArgLysLeuGlnGlyGluAsnGluGluGluGlyLysGlyAlaIleValThrValLysGly 259  
QY 841 GCCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGAGCTCCCGACGAAGAGAG 891  
Db 260 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu 279  
QY 892 GAATACGATGAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGAT 951  
Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294  
QY 952 AGAGGCGAGGGGAATCGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACATT 1011  
Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311  
QY 1012 GGTAGAAACAGATCCCTCGACATCTACAACCTCTCAAGCTGGTTCACCTCAAACTGCCAAC 1071  
Db 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331  
QY 1072 GATCTCAACCTCTTAATATCTAGGTGGCTTGGACCTAGTGTGCTGAATATGGAATCTCTAC 1131  
Db 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351  
QY 1132 AGGAATGCTTCTTCTCGCTCCTACACCAACCAACGACGACGACATCATATATCGATTG 1191  
Db 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371  
QY 1192 AGGGGACGGCTCACCTGCAAGTCGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAG 1251  
Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391  
QY 1252 CTTCAAGAGGTCACGTCTGTGTGTCGACAGAACTTCGCGCTCGCTGGAAGTCCAG 1311  
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411  
QY 1312 AGCGAACTTCGAATACGTGGCATTCAAGACAGACTCAAGGCCCGCAGCATAGCAACCTC 1371  
Db 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431  
QY 1372 GCGGTGAAAACTCCCGTCATAGATAACCTCCCGAGGAGGTGTTGCAATATTCATATGC 1431  
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnThrPheAsn 451  
QY 1432 CTCCAAGGAGGAGCAGCAAGCAGCTTAAGAACCAACACCCCTTCAAGTTCCTCGTT--- 1488  
Db 452 LeuArgGlnSerGlnValSerGluLeuLysTyrGluGlyAsnTrpGlyProLeuValAsn 471  
QY 1489 CCACCTCTCCAGCAG---TCTCCGAGG 1512  
Db 472 ProGluSerGlnGlnGlySerProArg 480

## RESULT 11

S14393  
legumin A2 precursor - fava bean  
C;Species: Vicia faba (fava bean)  
C;Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: S14393  
R;Schlesier, B.; Bassuener, R.; van Hai, N.; Muentz, K.  
Nucleic Acids Res. 18, 7146, 1990  
A;Title: The cDNA derived primary structure of two distinct legumin A subunit precursors  
A;Reference number: S14392; MUID:91088307; PMID:2263481

Mon Apr 11 05:25:05 2005

A;Accession: S14393  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-500 <SCH>  
A;Cross-references: UNIPROT:Q99304; EMBL:X55014; NID:g22007; PIDN:CAA38758.1; PID:g22008  
C;Superfamily: glycinin

Alignment Scores:  
Pred. No.: 3,29e-97 Length: 500  
Score: 1414.50 Matches: 280  
Percent Similarity: 67.30% Conservative: 76  
Best Local Similarity: 52.93% Mismatches: 100  
Query Match: 52.56% Indels: 73  
DB: 2 Gaps: 8

US-10-728-051-3 (1-1524) x S14393 (1-500)  
QY 1 CGGCAGCAACCGGAGGAGACGGTCCAGTCCAGCGCTCAATGCGCAGAGACTGAC 60  
DB 1414.50  
QY 23 ArgGluGlnSerGlnGlnAenGluCysGlnLeuGluArgLeuAspAlaLeuGluProAsp 42  
DB 67.30%  
QY 61 AATCGCATTGAATCAGAGGCGGTATCATTCAGACTTGGAAACCCCAACACAGAGGTTTC 120  
DB 52.93%  
QY 43 AsnArgileGluSerGluGlyGlyLeuileGluThrTriAsnProAsnAsnArgGlnPhe 62  
DB 52.56%  
QY 121 GAATCGCGCGCGTCCCTCTCTCGTCTAGTCTCGCGCAACGCCCTTCGTAGGCTT 180  
DB 2  
QY 63 ArgCysAlaSerValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgAPro 82  
DB 67.30%  
QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCAGCAAGAAAGGGGATATTTGGGTG 240  
DB 52.93%  
QY 83 TyrTyrSerAsnAlaProGlnGluileTyrileGlnGlnGlyAsnGlyTyrPheGlyMet 102  
DB 52.56%  
QY 241 ATATTCCTGTTGCTTCTAGACACTATGAAGAGCTCACACAGAGTCTCGATCTCAG 300  
DB 2  
QY 103 ValPheProSerCysProGluThrPheGluGluPro-----GlnGln 116  
QY 301 TCCCAAGACCAACCAAGAGCTCTCCAAGGAGAGACCAACCAAGCAACAGAGATAGT 360  
DB 67.30%  
QY 117 SerGlu-----GlnGlyGlu-----GlyGlyArgTyrArgAspSer 128  
QY 361 CACCAAGAGTGCACCGTTTCAGTGGGTGATCTCATTCAGTTCCTCCAGCGGTGCT 420  
DB 52.93%  
QY 129 HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleVal 148  
QY 421 TTCTGCTCTCAACGACACGACACTGATGTTGCTGTTCTTCTTCTTACTGACACCAAC 480  
DB 52.56%  
QY 149 PheTrpMetTyrAsnAspGlnAspThrProValIleAlaIleSerLeuThrAspIleGly 168  
QY 481 AACCAACGACCAACCGCTTATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACCGAG 540  
DB 67.30%  
QY 169 SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 188  
QY 541 CAAGAGTCTTAAAGTACAGCAACAAAGCAGACAAACAGCAGCAAGAGCTTACCATAT 600  
DB 52.93%  
QY 189 GlnGluPheLeuArgTyrGlnHisGln-----GlnGlyGlyLysGluGluGlnAspAsnAspGlyAsnAsnIlePhe 197  
QY 601 AGCCCATACAGCCCGCAAGTCCAGCTAGACAAAGAGAGCGTGAATTTAGCCCTCGAGGA 660  
DB 52.56%  
QY 197 -----GlnGlyGlyLysGluGluGlnAspAsnAspGlyAsnAsnIlePhe 720  
QY 661 CAGCAGCGCGCAGAGAACAGCAGGAGCAAGAAAGAAAGAAAGAGTGGAAACATCTTC 720  
DB 67.30%  
QY 198 -----GlnGlyGlyLysGluGluGlnAspAsnAspGlyAsnAsnIlePhe 712  
QY 721 AGCGGCTTCAGCGCGAGTTCCTGGAAACAAAGCTTCCAGTTCAGCAGACAGATAGTG 780  
DB 52.93%  
QY 213 SerGlyPheLysArgPheLeuGluAspAlaLeuAsnVal---AsnArgHisIleVal 231  
QY 781 CAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGCGCATTTGTGACGTAGAGGGA 840  
DB 52.56%  
QY 232 AspArgLeuGlnGlyArgAsnGlnGluAspGluGlyAlaIleValLysValLysGly 251

QY 841 GGCCTCAGATCTTGAGCCAGATAGAAAG-----AGAGTCCCGACGAA 885  
DB 1414.50  
QY 252 GlyLeuSerIleIleThrProGluArgGlnAlaArgHisProArgGlySerArgGln 271  
DB 67.30%  
QY 886 GAAGAGGAATACATGAAGATGATATGAA-----AGGCGTGGCAGGGGAAACAG 954  
DB 52.93%  
QY 272 GluGluAspGluAspGluAspGluGluGluGluGluGluGluGluGluGluGluGluGlu 291  
DB 52.56%  
QY 916 -----TACGATCAAGAGGATAG-----AGGCGTGGCAGGGGAAACAG 954  
DB 67.30%  
QY 292 ArgGlyGluAspGluAspAspLysGluLysArgHisSerGlnLysGlyGluSerArgArg 311  
QY 955 GGCAGGGGGAATGATATGAAAGAGAGATCTGACCGCAAGTCTAAAGAAACATCTGCT 1014  
DB 52.93%  
QY 312 HisGlyAspAsnGlyLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 331  
DB 52.56%  
QY 1015 AGAAGAGATCCCTGACATCTCAACCTCAAGCTGGTGTCTCACTCAAACTGCGCAACGAT 1074  
DB 67.30%  
QY 332 SerSerSerSerProAspIleTyrAsnProGlnAlaGlyArgIleLysThrValThrSer 351  
QY 1075 CTCAACCTTCTTAATCTTAGTGTGCTTGGACCTTAGTGTCTGAATATGAAATCTCTACAGG 1134  
DB 52.93%  
QY 352 LeuAspLeuProValLeuArgTyrLeuLysLeuSerAlaGluHisGlySerLeuArgLys 371  
DB 52.56%  
QY 1135 AATGATTTGTTGCTGCTCACTACCAACCAACGACACAGCATCATATATCGATTGAGG 1194  
DB 67.30%  
QY 372 AsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleLeuTyrAlaLeuLys 391  
QY 1195 GAGCGGCTCAGCTGCAAGTCTGCGACAGCAACCGCAACAGAGTGTACGAGGAGCTT 1254  
DB 52.93%  
QY 392 GlyArgAlaArgLeuGlnValValAsnValAsnGlyAsnThrValPheAspGluGluLeu 411  
QY 1255 CAAGAGGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314  
DB 52.56%  
QY 412 GluAlaGlyArgAlaLeuThrValProGlnAsnTyrAlaValAlaAlaLysSerLeuSer 431  
QY 1315 GAGAACTTCAATCTGCTGCTTCAAGACAGACTCAAGGCCAGCAGATAGCCCACTCGCC 1374  
DB 67.30%  
QY 432 AspArgPheThrTyrValAlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAla 451  
QY 1375 GGTGAAACTCGCTCATAGATACTGCTGCGAGGAGGTGGTGGCAATTCATATGGCTC 1434  
DB 52.93%  
QY 452 GlyThrSerSerValIleAsnAspMetProValAlaValAlaAlaThrPheAsnLeu 471  
QY 1435 CAAAGGAGCAGGAGCAGCTTAAAGAAACAAACCCCTTCAAGTCTTCGTTCCACCG 1494  
DB 52.56%  
QY 472 GluArgAsnGluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProPro 491  
QY 1495 TCTCAGCAGCTCTCCGAGGCTGTGCT 1521  
DB 67.30%  
QY 492 ArgGluSerGlnLysArgAlaSerAla 500

RESULT 12

T06452  
probable legumin A precursor - garden pea  
C;Species: Pisum sativum (garden pea)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-Jul-2000  
C;Accession: T06452  
R;Lycett, G.W.; Crox, R.D.; Shirsat, A.H.; Richards, D.M.; Boulter, D.  
Nucleic Acids Res. 13, 6733-6743, 1985  
A;Title: The 5'-flanking regions of three legumin genes: comparison of the DNA sequences  
A;Reference number: Z15687; MUID:86041868; PMID:2997721  
A;Accession: T06452  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-507 <LYC>  
A;Cross-references: EMBL:X02982; NID:g20777; PIDN:CAA26720.1; PID:g4379378  
C;Genetics:  
A;Introns: 96/1; 179/3; 388/3  
C;Superfamily: glycinin  
C;Keywords: seed; storage protein  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;122-332/Product: legumin A, alpha chain #status predicted <ALP>



DB:	2	Gaps:	11
US-10-728-051-3 (1-1524) x S14392 (1-497)			
QY	25	TGCCAGTTCACGGCTCAATCGCGAGACCTGACAATCGATTGAATCAGAGGGCGGT	84
Db	1	CysGlnLeuGluArgLeuAspAlaLeuGluProAspAsnArgIleGluSerGluGly	20
QY	85	TACATTGAGACTTGGAAACCCCAACACAGGAGTTCGAATGCGCGCGGTCTCTCT	144
Db	21	LeuIleGluThrTrpAsnProAsnAsnArgGlnPheArgCysAlaArgValAlaLeuSer	40
QY	145	CGTTAGTCTCCGCGCAACGCCCTTCGTAGCCCTTCTACTCCAATGCTCCCGAGAG	204
Db	41	ArgAlaThrLeuGlnArgAsnAlaLeuArgA-gProTyr-Tyr-SerAsnAlaProGlnGlu	60
QY	205	ATCTTCATCCAGCAAGGAGATCTTTGGGTTGATATTCCTGTTGTCTTAGACAC	264
Db	61	IleTyrIleGlnGlnArgGlnPheGlyMetValPheProSerCysProGluThr	80
QY	265	TATGAAGAGCTCACACAAAGTCTCGATCTCAGTCCCAAGACCAACCAAGAGTCTC	324
Db	81	PheGluGluPro-----GlnGlnSerGlu-----	88
QY	325	CAAGGAGAAGACCAAGCAACAGACGATAGTACCAGAAAGGTGCACCGTTTTCGAT	384
Db	89	GlnGlyGlu-----GlyArgArgTyrArgAspSerHisGlnLysValAsnArgPheArg	106
QY	385	GAGGGTGATCTCATTTGAGTTCACCGGTGTTGCTTTCTGGCTCTACAAACGACACGAC	444
Db	107	GlnGlyAspIleIleAlaValProThrGlyIleValPheTrpMetTyrAsnAspGlnAsp	126
QY	445	ACTGATGTTGCTGTTCTCTTACTGACCAACCAACCAACCAACCAAGCTTGTATCAG	504
Db	127	IleProValIleAlaIleSerLeuThrAspThrGlySerSerAsnAsnGlnLeuAspGln	146
QY	505	TTCCCCAGGAGATCAATTTGGCTGGGAACACGAGCAAGAGTCTTAAGGTACACGAA	564
Db	147	MetProArgArgPheTyrLeuAlaGlyAsnGlnGlnGluPheLeuArgTyrGlnHis	166
QY	565	CAAGAGCACAACGACGAGNAGAGTTCACCATATAGCCATACAGCCCGCAAGTCAG	624
Db	167	GlnGlnGlyValLysGluGlu-----GlnAspAsn	176
QY	625	CCTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGACGACGACGACGACGACGACGA	684
Db	177	AspGlyAsnGlnGlnGlnGluPheLeu---ArgTyrGlnHis-----ArgGlnGlyVal	193
QY	685	GGACAAGAGAGAAACGAAGGTGGAAACATCTTCAGCGGCTTCACCGCGGAGTTCCTG	744
Db	194	LysGluGluGlnAspAsnAspGlyAsnAsnIlePheSerGlyPheAsnArgAspPheLeu	213
QY	745	GNACAAGCTTCCAGGTTGACGACGACGACGATAGTGCMAAACCTTAAGAGGCGACCCGAG	804
Db	214	GluAspAlaPheAsnVal---AsnArgHisIleValAspArgLeuGlnGly-----Arg	230
QY	805	AGTGAAGAGAGGAGCCATTGTGACGTGAGGGAGGCCTCAGAACTTCGAGCCCGAT	864
Db	231	AsnGluGluArgGlyAlaIleValLysValLysGlyGlyLeuSerIleIleThrProPro	250
QY	865	AGAAGAGA-----CGTCCGACGAAGAAGAGGATACGATGAA	903
Db	251	GluArgGlnAlaArgHisProArgGlySerArgGlnGluGluAspGluAspGlu	270
QY	904	GATGAATATGAA-----TACGATGAGAG	927
Db	271	AspGluLysGluGluArgGlnProSerHisHisLysSerArgArgGlyGluAspGluAsp	290
QY	928	GATAGA-----AGGCGTGCAGGGGAACGACGAGGCGAGGGGGAATGTTATT	972
Db	291	AspLysGluLysArgHisSerGlnLysGlyGluSerArgArgHisGlyAspAsnGlyLeu	310
QY	973	GAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACATTTGGTGTAGAAACAGATCCCTCGAC	1032

Db	311	GluGluThrValCysThrAlaLysLeuArgLeuAsnIleGlySerSerSerProAsp	330
QY	1033	ATCTACAACCTCAAGCTGGTTCACTCAAACTGCCAACGATCTCAACTTCTTAATCTT	1092
Db	331	IleTyrAsnProGlnAlaGlyArgIleLysThrValThrSerValAspLeuProValLeu	350
QY	1093	AGTGGCTTGGACCTAGTCTGAATATGAAATCTCTACAGGAATCATTTGTTGCTCT	1152
Db	351	ArgTrpLeuLysLeuSerAlaGluHisGlySerLeuArgLysAsnAlaMetPheValPro	370
QY	1153	CATTACAACACCAACGACACACATCATATATCGATTGAGGGAGCGGCTCACGTGCAA	1212
Db	371	HisTyrAsnLeuAsnAlaAsnSerValLeuTyrAlaLeuLysGlyArgAlaArgLeuGln	390
QY	1213	GTCTGTGACAGCAACGCGACAGAGTGTACGACGAGAGCTTCAAGAGGGTCACTGCTT	1272
Db	391	ValValAsnCysAsnGlyAsnThrValPheAspGlyGluLeuGluAlaArgAlaLeu	410
QY	1273	GTGTGTCCACAGAACTTTCGCGTGTGGAAGTCCAGAGCGAGAACTTCGAATACGTG	1332
Db	411	ThrValProGlnAsnTyrValValAlaAlaLysSerLeuSerAspArgPheThrTyrVal	430
QY	1333	GCATTCAAGCAGACTCAAGCCGACATAGCCAACTCCCGGTGAAAACTCCGTCATA	1392
Db	431	AlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAlaGlyThrSerSerValIle	450
QY	1393	GATAACCTTCGCGAGGAGGTGGTTGCAATTCATATATGCGCTCCAAAGGAGGCGAAGG	1452
Db	451	AsnAspLeuProLeuAspValValAlaAlaThrPheAsnLeuGluArgAsnGluAlaArg	470
QY	1453	CAGCTTAAGAACACCAACCCCTTCAAGTCTTCTGTTCCACCGTCTCAG	1500
Db	471	GlnLeuLysPheAsnAsnProSerArgPheLeuValProProArgGlu	486
RESULT 14			
S51941			
prunin 1 precursor - almond			
N:Alternate names: salt-soluble globulin; seed storage protein			
C:Species: Prunus dulcis (almond)			
C>Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004			
C:Accession: S51941; S42474			
R:Garcia-Mas, J.; Messegue, R.; Arus, P.; Puigdomenech, P.			
Plant Mol. Biol. 27, 205-210, 1995			
A:Title: Molecular characterization of cDNAs corresponding to genes expressed during alc			
A:Reference number: S51940; MUID:95170003; PMID:7865791			
A:Accession: S51941			
A>Status: nucleic acid sequence not shown			
A:Molecule type: mRNA			
A:Residues: 1-551 <GAR>			
A:Cross-references: UNIPROT:Q43607; EMBL:X78119			
R:Garcia-Mas, J.; Messegue, R.; Arus, P.; Puigdomenech, P.			
submitted to the EMBL Data Library, March 1994			
A:Description: Molecular characterization of cDNAs corresponding to proteins related to			
A:Reference number: S42473			
A:Accession: S42474			
A:Molecule type: mRNA			
A:Residues: 1-60, G' 62-551 <GAW>			
A:Cross-references: EMBL:X78119; NID:g460805; PIDN:CAA55009.1; PID:g460806			
A>Note: the source is designated as Prunus amygdalus			
C:Superfamily: glycinin			
C:Keywords: seed; storage protein			
F:1-20/Domain: signal sequence #status predicted <SIG>			
F:21-367/Product: prunin 1 alpha chain #status predicted <MAT1>			
F:368-551/Product: prunin 1 beta chain #status predicted <MAT2>			
F:108-374/Disulfide bonds: #status predicted			
Alignment Scores:			
Pred. No.:	7,47e-72	Length:	551
Score:	1073.50	Matches:	228
Percent Similarity:	57.22%	Conservative:	89
Best Local Similarity:	41.16%	Mismatches:	160



Query Match: 39.89% Indels: 77  
 DB: 2 Gaps: 11  
 US-10-728-051-3 (1-1524) x SS1941 (1-551)

QY 4 CAGCAACGGAGGAGAGCGGTTCAGCGCTCAATGCGCAGACCTGCAAT 63  
 DB 25 GlnLeuSerProGlnAsnGlnCysGlnLeuAsnGlnLeuGlnAlaArgGluProAspAsn 44  
 QY 64 CGCATTAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTGAA 123  
 DB 45 ArgIleGlnAlaGluAlaGlyGlnIleGluThrTrpAsnPheAsnGlnGluAspPheGln 64  
 QY 124 TGGCGCGGCTCGCCCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGCCCTTC 183  
 DB 65 CysAlaGlyValAlaAlaSerArgIleThrIleGlnArgAsnGlyLeuHisLeuProSer 84  
 QY 184 TACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATATCTTTGGTTGATA 243  
 DB 85 TyrSerAsnAlaProGlnLeuIleIleValGlnArgGlyValLeuGlyAlaVal 104  
 QY 244 TTCCCTGGTTGCTCAGACACTATGAAGAG----- 273  
 DB 105 PheSerGlyCysProGluThrPheGluGluSerGlnGlnSerSerGlnGlnArgGln 124  
 QY 274 ---CCTCACACACAGTCTCGATCTCAGTCCCAAGACCAACAGACGCTCTCCAAGGA 330  
 DB 125 GlnGluGlnGluGlnGluArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 144  
 QY 331 GAAGACCAAAACCAACAGCAGCAAG----- 354  
 DB 145 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 164  
 QY 354 ----- 354  
 DB 165 GlyArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnPhe 184  
 QY 355 -----GATAGTCACAGAGGTGCACCGTTTCATGAGGTGATCTCATTCAGTT 405  
 DB 185 ArgGlnLeuAspArgHisGlnLysThrArgArgIleArgGluGlyAspValAlaIle 204  
 QY 406 CCAACCGGTGTGCTTCTCGCTCTACACGACACCACTGATGTTGTTGTTCT 465  
 DB 205 ProAlaGlyValAlaIleTyrTrpSerTyrAsnAspGlyAspGlnGluLeuValAlaValAsn 224  
 QY 466 CTTACTGACCAACCAACACAGCAGCTGATCAGTTCCTCCAGGAGATTCATTTG 525  
 DB 225 LeuPheHisValSerSerAspHisAsnGlnLeuAspGlnAsnProArgLysPheTyrLeu 244  
 QY 526 GTGGGAAACACGGAGCAAGAGTTCTTAAGGTACAGCAACAAAGCAGACAAACGACGA 585  
 DB 245 AlaGlyAsnProGluAsnGluPhe-----AsnGlnGlnGly----- 256  
 QY 586 AGAAGCTTACCATATACCCCATACAGCCCGCAAGTCAGCTAGACAAAGACGCTGAA 645  
 DB 257 -----GlnSerGlnProArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 266  
 QY 646 TTTAGCCCTCGAGCAGCAGCAGCCGC-----AGAAACGAGCAGCAGCAAGAA 696  
 DB 267 GlyArgPro---GlyGlnHisGlnGlnProPheGlyArgProArgGln---GlnGlnGln 284  
 QY 697 GAAACGAAGGTGAAACATCTTCAGCGGTTTCACGCGGAGTTCCTGGAACAGGCTTC 756  
 DB 285 GlnGlyAsnGlyAsnAsnValPheSerGlyPheAsnThrGlnLeuLeuAlaGlnAlaLeu 304  
 QY 757 CAGTTACGACACACAGATAGTCAAACTAGACCGCAGACCGAGAGTGAAGAGAG 816  
 DB 305 AsnValAsnGlu---GluThrAlaArgAsnLeuGlnGlnGlnAsnAsnArgAsnGln 323  
 QY 817 GGAGCCATTGTGACAGTGGGGAGGCGCTCAGAAATCTTGAAGCCAGATAGAAGAGAGT 876  
 DB 324 -----IleIleGlnValArgGlyAsnLeuAspPheValGlnProProArgGlyArgGln 341

877 GCCGACGAAGAGAGGAATACGATGAAGATGAATATGAATACGATGAAGAGGATAGAGG 936  
 DB 342 GluArgGluHisGluGluArgGlnGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGln 357  
 QY 937 CGTGGCAGGGAAGAGCAGAGCGGGGAATGTGATTGAAGAGAGCATCTCACCAGCAAGT 996  
 DB 358 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 377  
 QY 997 GCTAAAAAGAACATTCGTAGAAAACAGATCCCTGACATCTCAACCTCAAGCTGGTTCA 1056  
 DB 378 LeuLysGluAsnIleGlyAsnProGluArgAlaAspIlePheSerProArgAlaGlyArg 397  
 QY 1057 CTCAAACTCCCAACCATCTCAACCTCTAATATTAGTGGCTTGGACCTAGCTAGTGA 1116  
 DB 398 IleSerThrLeuAsnSerHisAsnLeuProIleLeuArgPheLeuArgLeuSerAlaGlu 417  
 QY 1117 TATGGAATCTCTACAGGAATGATTGTTGTCTGCTCACTACCAACACCAACGACACAGC 1176  
 DB 418 ArgGlyPhePheThrArgAsnGlyIleTyrSerProHisTrpAsnValAsnAlaHisSer 437  
 QY 1177 ATCATATATGATTGAGGGACGGGCTCAGCTCAAGTCTGTCAGCAGCAACGCAACAGA 1236  
 DB 438 ValValTyrValIleArgGlyAsnAlaArgValGlnValValAsnGluAsnGlyAspAla 457  
 QY 1237 GTGTACAGCAGGAGCTTCAAGAGGCTCAGCTGCTGTGTGTGTGTGTGTGTGTGTGT 1296  
 DB 458 IleLeuAspGlnGluValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 477  
 QY 1297 GCTGGAAGTCCACAGAGCAGAACTTCGAATACGTGCTCACTTCAAGACAGACTCAAG 1356  
 DB 478 IleGlnGlnAlaGlyAsnGlnGlnGlyPheGluTyrPheAlaPheLysThrGluGluAsnAla 497  
 QY 1357 AGCATAGCCAACTCCCGGTGAAACTCCGTCTAGATAACCTCCGCGGAGGAGTGTTG 1416  
 DB 498 PheIleAsnThrLeuAlaGlyArgThrSerPheLeuArgAlaLeuProAspGluValLeu 517  
 QY 1417 GCAAATTCATATGCTCCAAAGGAGCAGGCAAGCAGCTTAAGCAACAACAACCCCTTC 1476  
 DB 518 AlaAsnAlaTyrGlnIleSerArgGluGlnAlaArgGlnLeuLysTyrAsnArgGlnGlu 537  
 QY 1477 AGTTCTTCTGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTG 1518  
 DB 538 ThrIleAlaLeuSerSerSerGlnGlnArg---ArgAlaVal 550

RESULT 15  
 S11004  
 C:Species: glycinin G4 precursor - soybean  
 C:Species: Glycine max (soybean)  
 C:Date: 21-Nov-1993 #sequence\_revision 19-Jan-1996 #text\_change 09-Jul-2004  
 C:Accession: S11004  
 R:Nielson, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallan, B.J.; Fischer, R.L.;  
 Plant Cell 1, 313-328, 1989  
 A:Title: Characterization of the glycinin gene family in soybean.  
 A:Reference number: S10851; MUID:92393391; PMID:2485233  
 A:Accession: S11004  
 A:Molecule type: DNA  
 A:Residues: 1-560 <NIE>  
 A:Cross-references: UNIPROT:Q9S9D0  
 A:Experimental source: variety Dare  
 C:Genetics:  
 A:Gene: GY4  
 C:Superfamily: glycinin  
 C:Keywords: storage protein  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-560/Product: glycinin G4 #status predicted <MAT>

Alignment Scores:  
 Pred. No.: 4,96e-70 Length: 560  
 Score: 1049.00 Matches: 227  
 Percent Similarity: 53.33% Conservative: 77  
 Best Local Similarity: 39.82% Mismatches: 158  
 Query Match: 38.98% Indels: 108  
 DB: 2 Gaps: 10



334	AspGlnAspGlnAspGluAspGluAspGluAspGlnProArgLysSerArgGlu	353
958	-----AGGGGG	963
354	TrpArgSerLysLysThrGlnProArgProArgGlnGlnGluProArgGluArgGly	373
964	-----AATGGTATTGAAGAGACGATCTGCACCGCAAGTCTTAAAAAGACATT	1011
374	CysGluThrArgAsnGlyValGluGluAsnIleCysThrLeuLysLysLeuHisGluAsnIle	393
1012	GGTAGAAACAGATCCCTCGACATCTCAACCCCTCAAGCTGGTTTCATCTAAAACTGCCAAC	1071
394	AlaArgProSerArgAlaAspPheTyAsnProLysAlaGlyArgLysThrLeuAsn	413
1072	GATCTCAACCTTCTTAATCTAGTGCGCTGGACCTAGTCTGCTGAATATGAAATCTCTAC	1131
414	SerLeuThrLeuProAlaLeuLeuArgGlnPheGlnLeuSerAlaGlnTyValValLeuTy	433
1132	AGGATCGATTGTTGCTGCTCACTCAACACCAACGACACACAGCATCATATATCGATTG	1191
434	LysAsnGlyIleTySerProHisTrpAsnLeuAsnAlaAsnSerValIleTyValThr	453
1192	AGGGCAGCGGCTCACGTCGAAGTCGTGGACAGCAACGCGCAACAGAGTGTACGACGAGG	1251
454	ArgGlyGlnGlyLysValArgValValAsnCysGlnGlyAsnAlaValPheAspGlyGlu	473
1252	CTTCAAGAGGGTACAGTCTGTGTGGTCCACAGAACTTCGCCGCTCGTGGAAAGTCCAC	1311
474	LeuArgGlyGlnLeuLeuValValProGlnAsnPheValValAlaGluGlnAlaGly	493
1312	AGCCAGAACTTCGAATACGTGGCATTCAGACACAGACTCAAGGCCCGCAGCATAGCCAACTC	1371
494	GluGlnGlyPheGluTyIleValPheLysThrHisAsnAlaValThrSerTyLeu	513
1372	GCCGGTGAATACTCCGTATAGATAAATCGCCGAGGAGGTGGTTCGAAATTCATATGCG	1431
514	-----LysAspValPheArgAlaIleProSerGluValLeuAlaHisSerTyAsn	530
1432	CTCCAAAGGGAGCAGCGCAAGCCAGTCTAAGAACCAACACCCCTTCAGTTCCTCGTCCCA	1491
531	LeuArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro	550
1492	CCGTCTCAGCAGTCTCCGAGGGCTGTGGCT	1521
551	ProLysGluSerGlnArgArgValValAla	560

Search completed: April 8, 2005, 05:30:59  
 .Job time : 90.7154 secs

US-10-728-051-3 (1-1524) x S11004 (1-560)

QY	19	AACGCGTGCAGTTCCAGCGCCTCAATGCGCAGAGACCTGACAATCGCATTTGAATCAGAG	78
DB	30	AsnGluCysGlnLeuAenAenLeuAenAlaLeuGluProAspHisArgValGluSerGlu	49
QY	79	GGCGGTACATTGAGACTTCGGAACCCCAACACACAGGAGTTCGAATGCGCGCGCTCGCC	138
DB	50	GlyGlyLeuIleGlnThrTrpAsnSerGlnHisProGluLeuLysCysAlaGlyValThr	69
QY	139	CTCTCTCGCTTAGTCTCCCGCAACGCCCTTCGTAGGCGTTTCTACTCCAAANGTCTCC	198
DB	70	ValSerLysLeuThrLeuAenArgAsnGlyLeuHisLeuProSerTrpSerProTrpPro	89
QY	199	CAGAGAGATCTTCACGACGAGGAGGAGATCTTTGGTGTGATATCTCCCTGGTGTCT	258
DB	90	ArgMetIleIleAlaGlnGlyLysValAlaLeuGlyValAlaIleProGlyCysPro	109
QY	259	AGACACTATGAAGACGCTCACACACAAAGTCGTGATCTCAGTCCCAACACCAACCAAGA	318
DB	110	GluThrPheGluGluProGlnGlnSerAsnArgArgGlySerArg	125
QY	319	CGTCTCCAAGGAGAACCAAGACACACAGCATAGTCTACGAGAAGTGCACCGT	378
DB	126	-----SerGlnLysGlnLeuGlnAspSerHisGlnLysIleArgHis	140
QY	379	TTGATGAGGGTGATCTCATTTGCAGTTCCCAACCGGTGTGCTTCTCGCTCTACAAACGAC	438
DB	141	PheAsnGluGlyAspValLeuValIleProGlyValProTrpThrTrpThrValThr	160
QY	439	CACGACACTGATGTTGTTGCTGTTCTCTACTGACCAACACACAGCAACACAGCTT	498
DB	161	GlyAspGluProValValAlaIleSerLeuLeuAspThrSerAsnPheAsnAsnGlnLeu	180
QY	499	GATCATGTTCCCGAGAGATTCAATTGGCTGGGAACGCGAGAGAGTCTTTAAGGTAC	558
DB	181	AspGlnThrProArgValPheTrpLeuAlaGlyAsnProAspIleGluTrp	197
QY	559	CAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATTATAGCCCATACAGCCGCA	618
DB	198	-----ProGluThrMetGln	202
QY	619	AGTCGCTTAGACAAGAAGAGCGTGAAATTAGCTTCGAGACACGACGCGCAGAGAA	678
DB	203	GlnGlnGlnGlnLysSerHisGlyArgLysGlnGlyGlnHisGln	219
QY	679	CGAGCAGGACAGAGAGAGAAACCAAGGTGAAACATCTTCAGCGGTTCACGCGCGAG	738
DB	220	-----GlnGluGluGluGluGlySerValLeuSerGlyPheSerLysHis	236
QY	739	TTCTCTGGAAACAAGCCTTCAGTTGACGACACAGATAGTGTCAAAACCTAAGGCGCAG	798
DB	237	PheLeuAlaGlnSerPheAsnThrAsnGlu-----AspIleAlaGluLysLeu-----Gln	253
QY	799	ACCGAGGTGAAGACAGGAGGCCATTGTACAGTGTGCGGAGGCGCTCAGATCTTGAGC	858
DB	254	SerProAspGluArgLysGlnIleValThrValGluGlyGlyLeuSerValIleSer	273
QY	859	CCAGATAGAAAGACGCTGCCACGAAGAGAGGAATACGATCAA-----	903
DB	274	ProLysTrpGlnGlnGlnAspGluAspGluAspGluAspGluAspGluAspGlu	293
QY	904	-----CATGAATAT	912
DB	294	GlnIleProSerHisProProArgProSerHisGlyLysArgGlnAspGluAsp	313
QY	913	GAATACGATCAACAGGATAGA-----AGCGGTGGCAGGGGAAGCAGAGCC-----	957
DB	314	GluAspGluAspGluAspLysProArgProSerArgProSerGlnGlyLysArgGluGln	333
QY	957	-----	957



52 AGACCTGACAAATCGCAATTCGAATTCAGAGGCGGTTTACATTGACACTTGGAAACCCCAAC 111  
45 GluProSerPheArgPheProSerGluAlaGlyLeuThrGluPheTrpAspSerAsn 64  
112 CAGGAGTTCGAATCGCGCGCTCGCCCTCTCGCTTAGTCTCCGCGCGCAACCCCTT 171  
65 ProGluPheGlyCysAlaGlyValGluPheGluArgAsnThrValGlnProLysLeu 84  
172 CGTAGGCTTTCCTTCTCCATGCTCCCAAGGAGATCTTCATCCAGCAAGAGGAGTAC 231  
85 ArgLeuProHisTyrSerAsnValProLysPheValTyrValValGlyThrGlyVal 104  
232 TTTGGGTGATATCCCTGTTCTCTAGACACTATGAGAGCCTCACACAAAGTCTGT 291  
105 GlnGlyThrValIleProGlyCysAlaGluThrPheGlu-----SerGlnGlyGlu 121  
292 CGA-----TCTCAGTCCCAAGACCAACCAAGCGTCTCCAAGGAGAACACCA 339  
122 SerPheTrpGlyGlnGluGlnProGlyLysGlyGlnGluGlnGlnGlySer 141  
340 -----AGCCCAACAGCAACGA-----GATAGTACACAGAGGTGACCGT 378  
142 LysGlyGlnGluGlyArgGlnArgPheProAspArgHisGlnLysLeuArg 161  
379 TTCATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTTCTTCTGGCTCTACAAAC 438  
162 PheGlnLysGlyAspValLeuLeuLeuLeuProGlyPheThrGlnTrpTyrAsnAsp 181  
439 CACGACTGATGTGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 498  
182 GlyAspValProLeuValThrValAlaLeuLeuAspValAlaAsnGluAlaAsnGlnLeu 201  
499 GATCAGTTCCTCCAGAGATTCATTCCTGCTGGGACACGGAGCAAGTCTTAAAGTAC 558  
202 AspLeuGlnSerLysPhePheLeuAlaGlyAsnProGlnGlnGlyGlyLysGlu 221  
559 CAGCAACAAAGCAGCAAGACGAGCAAGAGCTTACCATATAGCCCATACAGCCCGCAA 618  
222 GlyHisGlnGlnGlnGlnGlnHisArg----- 231  
619 AGTCAGCTAGCAAGAGCGGTGAATTTAGCCCTCGAGGACAGCAGCCGCGAGAA 678  
231 ----- 231  
679 CGAGCAGCAGAGAGAGAAACGAGGTGGAACATCTTCAGCGGCTTCACCGCGAG 738  
232 -----AsnIlePheSerGlyPheAspGln 240  
739 TTTCTGGAACAGCCTTCAGGTTGACGACAGACATAGTGCAAAACCTTAGAGCGAG 798  
241 LeuLeuAlaAspAlaPheAsnVal---AspLeuLysIleLeuGlnLysLeuLysGly--- 258  
799 ACCGAGGTGAGAGAGAGGAGCCATTGTGACAGTGAGGGA---GGCCTCAGAACTTTG 855  
259 ---ProLysAspGlnArgGlySer-----ThrValArgAlaGluLysLeuGlnLeuPhe 275  
856 AGCCCAAGTAGAAGAGAGCTCCGACGAGAGAGAGATACGATGAAGATGAATAGAA 915  
276 LeuProGluTyrSerGluGlnValGlnGlnProGlnGln----- 288  
916 TAGCATGAGAGATAGAGCGGTGCGAGGAGAGCAGGAGGAGGAGGAGTATGTTGAA 975  
289 ---GlnGlnGlnGlnGlnHisGlyValGlyArgGlyTrpArgSerAsnGlyLeuGlu 307  
976 GAGCAGATCTGACCGCAAGTGTAAAGAAACATTTGTAGAAACAGATCCCTCGACATC 1035  
308 GluThrLeuCysThrValLysLeuSerGluAsnIleGlyLeuProGlnGluAlaAspVal 327  
1036 TACAACCTCAAGCTGTTCACTCAAACTGCCACGATCTCAACCTTCTATCTAGG 1095  
328 PheAsnProArgAlaGlyArgIleThrValAsnSerGlnLysIleProLysLeuSer 347  
1096 TGGCTTGGACCTAGTCTGAATATGGAATCTCTACAGGAATGCAATTTTCTGCTCAC 1155

348 SerLeuGlnLeuSerAlaGluArgGlyPheLeuTyrSerAsnAlaIlePheAlaProHis 367  
1156 TACAACACCAACGACACACAGCATCATATATTCATTGAGGCGACGGCTCACGTGCAAGTC 1215  
368 TrpAsnIleAsnAlaHisAsnAlaLeuTyrValIleArgGlyAsnAlaArgIleGlnVal 387  
1216 GTGCACAGCAACCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGCTCACTGCTTGTG 1275  
388 ValAspHisLysGlyAsnLysValPheAspAspGluValLysGlnGlyLeuIleIle 407  
1276 GTGCCACAGAACTTCGCTCGCTGGAAGTCCACAGAGCGAGAACTTCGAATACGTGGCA 1335  
408 ValProGlnTyrPheAlaValIleLysLysAlaGlyAsnGlnGlyPheGluTyrValAla 427  
1336 TTCAGACAGCACTCAAGCGCCAGCATAGCCCACTCGCGGTGAAACCTCCGTCATAGAT 1395  
428 PheLysThrAsnAspAsnAlaMetIleAsnProLeuValGlyArgLeuSerAlaPheArg 447  
1396 AACCTGCGGAGGAGGTGTGCAATTCATATGCTCCAAAGGAGCAGGCAAGGCAG 1455  
448 AlaIleProGluGluValLeuArgSerSerPheGlnIleSerGluGluAlaGluGlu 467  
1456 CTTAAGACACACACCCCTTCAAGTCTTCTTCCACCGTCTCAGCAGTCT 1506  
468 LeuLysTyrGlyArgGlnGluArgLeuLeu-----SerGluGlnSer 482

RESULT 2  
US-09-645-593-12  
; Sequence 12, Application US/09645593  
; Patent No. 6777591  
; GENERAL INFORMATION:  
; APPLICANT: Chaudhary, Sarita  
; APPLICANT: van Rooijen, Gifs  
; APPLICANT: Moloney, Maurice  
; APPLICANT: Singh, Surinder  
; TITLE OF INVENTION: Flax Seed Specific Promoters  
; FILE REFERENCE: 9369-151  
; CURRENT APPLICATION NUMBER: US/09/645,593  
; CURRENT FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: US 60/151,044  
; PRIOR FILING DATE: 1999-08-27  
; PRIOR APPLICATION NUMBER: US 60/161,722  
; PRIOR FILING DATE: 1999-10-27  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Linum usitatissimum  
US-09-645-593-12

Alignment Scores:  
Pred. No.: 1,21e-19 Length: 141  
Score: 293.00 Matches: 55  
Percent Similarity: 72.73% Conservative: 25  
Best Local Similarity: 50.00% Mismatches: 30  
Query Match: 10.89% Indels: 0  
DB: 4 Gaps: 0

US-10-728-051-3 (1-1524) x US-09-645-593-12 (1-141)

QY 1138 GCATTGTTTGTCTCTCACTACACCAACGACATCATATATATCATGTTAGGGCA 1197  
Db 2 AlaIleArgLeuProHisTrpAsnIleAsnAlaHisSerIleValTyrAlaIleArgGly 21  
QY 1198 CGGCTCAGCTGCAAGTCTGGAGCAACGCGCAACAGAGTGTACACGAGGAGCTTCAA 1257  
Db 22 GlnAlaArgValGlnIleValAsnGluGluGlyAsnSerValPheAspGlyValLeuGln 41  
QY 1258 GAGGCTCAGCTGCTGTTGGTGGCCACAGAACTTCGCCCTCGCTCGAAAGTCCCGACGAG 1317  
Db 42 GluGlyGlnValValThrValProGlnAsnPheAlaValValLysArgSerGlnSerGlu 61

```
QY 1318 AACTTCGAATACGTGGCAATTCAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGT 1377
Db 62 ArgPheGluTrpValAlaPheLysThrAsnAspAlaMetValAsnSerLeuAlaGly 81
QY 1378 GAAAACTCCGTCTAGATAACCTGCCGAGAGGTGGTTCGCAAAATTCATATGCGCTCCAA 1437
Db 82 ArgThrSerAlaValArgAlaIleProAlaaspValLeuAlaAsnAlaIleTrpArgValSer 101
QY 1438 AGGAGCAGGCAAGCGCAGCTTAAGACAAC 1467
Db 102 ProGluGluAlaArgArgValLysPheAsn 111

RESULT 3
US-09-645-593-10
; Sequence 10, Application US/09645593
; Patent No. 6777591
; GENERAL INFORMATION:
; APPLICANT: Chaudhary, Sarita
; APPLICANT: van Rooijen, Gijb
; APPLICANT: Moloney, Maurice
; APPLICANT: Singh, Surinder
; TITLE OF INVENTION: Flax Seed Specific Promoters
; FILE REFERENCE: 9369-151
; CURRENT APPLICATION NUMBER: US/09/645,593
; CURRENT FILING DATE: 2000-08-25
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: US 60/151,044
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Linum usitatissimum
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: n is any amino acid
; NAME/KEY: SITE
; LOCATION: (62)
; OTHER INFORMATION: n is any amino acid
; NAME/KEY: SITE
; LOCATION: (78)
; OTHER INFORMATION: n is any amino acid
; NAME/KEY: SITE
; LOCATION: (82)
; OTHER INFORMATION: n is any amino acid
US-09-645-593-10

Alignment Scores:
Pred. No.: 1,98-13 Length: 85
Score: 227.50 Matches: 46
Percent Similarity: 61.22% Conservative: 14
Best Local Similarity: 46.94% Mismatches: 25
Query Match: 8.45% Indels: 13
DB: Gaps: 2

US-10-728-051-3 (1-1524) x US-09-645-593-10 (1-85)
QY 220 GGAAGGGATACCTTGGTGTGATATCCCTGGTGTCTTACACATATGAGAGCCTCAC 279
Db 1 GlyArgGlyValThrGlyIleMetPhePro***CysProGluThrPheGluGluSerGln 20
QY 280 ACACAAGGTGCTGCATCTCAGTCCCAAGACCACCAAGACGCTCTCCAGGAGAGACCAA 339
Db 21 GlnGlnGlyGlnGlnGlyGln-----GlnGlySerSerGln 32
QY 340 AGCCAACAGCAACAGAGATAGTCACAGAGAGGTGCACCGTTTCGATGAGGGTGTCTCAT 399
Db 33 -----AspGlnHisGlnIleValIleArgArgPheArgGluGlyAspValIle 47
```

```
QY 400 GCAGTTCCACCGGTCTTCTGGCTCTACAACGACCACGACACTGATGTGTGCT 459
Db 48 AlaValProAlaGlyValAlaHisTrpSerTyrAsnAspGlyAsnGluProValMetAla 67
QY 460 GTTCTCTTACTCACACCAACAACAACGACACAGCTTTCATCAGTTCCCCAGG 513
Db 68 IleValValHisAspThrSerSerHisLeuAsnGlnLeuAspAsnProArg 85

RESULT 4
US-09-645-593-11
; Sequence 11, Application US/09645593
; Patent No. 6777591
; GENERAL INFORMATION:
; APPLICANT: Chaudhary, Sarita
; APPLICANT: van Rooijen, Gijb
; APPLICANT: Moloney, Maurice
; APPLICANT: Singh, Surinder
; TITLE OF INVENTION: Flax Seed Specific Promoters
; FILE REFERENCE: 9369-151
; CURRENT APPLICATION NUMBER: US/09/645,593
; CURRENT FILING DATE: 2000-08-25
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: US 60/151,044
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: US 60/161,722
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Linum usitatissimum
US-09-645-593-11

Alignment Scores:
Pred. No.: 4,298-12 Length: 165
Score: 215.00 Matches: 62
Percent Similarity: 43.06% Conservative: 28
Best Local Similarity: 29.67% Mismatches: 69
Query Match: 7.99% Indels: 50
DB: Gaps: 6

US-10-728-051-3 (1-1524) x US-09-645-593-11 (1-165)
QY 517 TTCAATTTGGCTGGAAACGAGAGAGTTCTTAAAGTACCAGCAACAAGACAGACAA 576
Db 2 PheTyrLeuAlaGlyAsnProArgAspGluPheGluGlnSerGlnGlnGlyArgLeu 21
QY 577 AGCAGACGAGAGAGCTTACCATATAGCCCATACAGCCCGCAAGTACGCTAGACAAGA 636
Db 22 SerArgGlyGluSer----- 26
QY 637 GAGCGTGAATTTAGCCCTCGAGGACACACAGCCGCGAGAGAACCGAGCAGGACAAGAAG 696
Db 27 -----GluGlyGlyArgGlyArgGluProLeuGlnProAlaThr 40
QY 697 GAAAACGAGGTGGAACAATCTTCAGCGGTTCACGCCGAGTTCCTGGAAACAGCCTTC 756
Db 41 ThrSerSer-----CysGlyIleAspSerLysLeuIleAlaGluAlaPhe 55
QY 757 CAGGTTCAACGACAGACAGATAGTGCAGAACCTTAAGCGCCGAGACCGAGAGTGAAGAAG 816
Db 56 AsnValAspGlu---AsnValAlaAlaArgArgLeu-----GlnSerGluAsnAspAsnArg 72
QY 817 GGAGCCATTTGTGACGTAGGGGGAGCGCTCAGAAATCTTGAGCCCAAGATAGAAAGAGACGT 876
Db 73 GlyGlnIleValArgValGluGlyGluLeuAspIleValArgProProThrSerIleGln 92
QY 877 GCCGACGAAAGAGGAATACGATGAATATGAATACGATGAAGAGGATAGAGG 936
Db 93 GluGluSerGlnGluGln----- 98
QY 937 CTGGCGAGGGGAGCAGGAGCGAGGGGG-----AATGTTATTGAGACGATC 984
```

Db 99 -----GlycylArgGlyGlyArgTyTyrSerAsnGlyValGluGluThrPhe 115  
QY 985 TCACCGCAAGTCCTAAAAGACATTGGTAGAACAGATCCCTGACATCTACAAACCT 1044  
Db 116 CysSerMetArgLeuLeuGluAsnIleGlyAspProSerArgAlaAspIlePheThrPro 135  
QY 1045 CAAGCTGGTTCACCTCAAACTGCCACGATCTCAACCTTCTTAATACTTAGGTGGCTTGA 1104  
Db 136 GluAlaGlyArgValArgSerLeuAsnSerHisAsnLeuProValLeuGlnTrpIleGln 155  
QY 1105 CTTAGTGTGAATATGAAATCTCTAC 1131  
Db 156 LeuSerAlaGluArgGlyValLeuTyr 164

RESULT 5  
US-09-424-283-3  
; Sequence 3, Application US/09424283  
; Patent No. 6437219  
; GENERAL INFORMATION:  
; APPLICANT: Grimes, et al.  
; TITLE OF INVENTION: Sucrose binding proteins  
; FILE REFERENCE: 4630-50206  
; CURRENT APPLICATION NUMBER: US/09/424,283  
; CURRENT FILING DATE: 1999-11-19  
; PRIOR APPLICATION NUMBER: PCT/US98/10465  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: US 60/047,568  
; PRIOR FILING DATE: 1997-05-22  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 489  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-424-283-3

Alignment Scores:  
Pred. No.: 6.51e-09 Length: 489  
Score: 184.50 Matches: 99  
Percent Similarity: 34.19% Conservative: 87  
Best Local Similarity: 18.20% Mismatches: 187  
Query Match: 6.86% Indels: 171  
DB: 4 Gaps: 17

US-10-728-051-3 (1-1524) x US-09-424-283-3 (1-489)

QY 1 CGGCAGCAACGGAGGAGACGGTGCAGTTCAGCGCCTCAATGCGGACAGACCT--- 57  
Db 80 ArgGluLysGluGluGluHisGlnGluGlnHisGluGluGluGluGluAsnProTyr 99  
QY 58 -----GACAATCGCATTGAATCAGAGCGCGGTACATTGAG 93  
Db 100 ValPheGluGluAspLysAspPheSerThrArgValGluThrGluGlySerIleArg 119  
QY 94 ACTTGGAAACCCCAACACCGAGGTTCGAA---TGCGCGCGCGTGCCTCTCTCGCTTA 150  
Db 120 ValLeuLysLysPheThrGluLysSerLysLeuLeuGlnGlyIleGluAsnPheArgLeu 139  
QY 151 GTCCTC-----CGCGCAACGCCCTTCGTAGGCGCTTCTACTCCAATGCTCCCCAG 201  
Db 140 AlaIleLeuGluAlaArgAlaHisThrPheValSerProArgHisPheAspSerGluVal 159  
QY 202 GAGATCTTCATCCAGCAAGGAGGGGATCTTTGGGTGGATATTCCTCGTGTGCTCCTAGA 261  
Db 160 ValLeuPheAsnIleLysGlyArgAlaValLeuGlyLeuVal----- 173  
QY 262 CACTATGAAGAGCCTCACACACAGGTGCTGATCTCAGTCCCAACACCAACCAAGACGT 321  
Db 173 ----- 173  
QY 322 CTCCAAGGAGAGACCAAGCCACACGACGATAGTCACCAAGAGGTGCACCGTTTC 381  
Db 174 -----ArgGluSerGluThrGluLysIleThrLeu 183

QY 382 GATGAGGCTGATCTCATTGCTCCACCGGTGTGCTTTCTGGCTCTACAACGACAC 441  
Db 184 GluProGlyAspMetIleHisIleProAlaGlyThrProLeuTyrIleValAsnArgAsp 203  
QY 442 GACACTGATGTTGTTGCTGTTTCTTACTGACACCAACAACAGCACACCGCTTGTAT 501  
Db 204 GluAsnGluLysLeuLeuAlaMetLeu----- 213  
QY 502 CAGTTCCCGCAGGAGATTCAATTGGCTGGGAACACGGAGCAAGAGTCTTAAAGTACCAG 561  
Db 213 ----- 213  
QY 562 CAACAAAGCAGCAAAAGCAGACGAGAAGCTTACCATATAGCCCATACAGCCCGCAAGT 621  
Db 214 -----HisIleProValSer 218  
QY 622 CAGCCTAGACAAGAAGCGGTGAATTTAGCCCTCGAGGACACGACGCGGACAGAACGA 681  
Db 219 ThrProGlyLysPheGluGluPheGlyProGlyGlyArg----- 232  
QY 682 GCAGGACAAGAAGAAGAAACGAAGGTGGAACATCTTCAGCGGCTTCACGCGGAGTTC 741  
Db 233 -----AspProGluSerValLeuSerAlaPheSerTrpAsnVal 245  
QY 742 CTGGAACAAGCCTTCAGGTTGACGACACAGATAGTGCACAAACCTAAGAGCGGACGCC 801  
Db 246 LeuGlnAlaLeuGlnThrPro-----LysGlyLysLeu 257  
QY 802 GAG-----AGTGAAGAAGGAGGAGCCATTGTGACAGTG--AGGGAGGCGCTC 846  
Db 258 GluArgLeuPheAsnGlnGlnAsnGluGlySerIlePheLysIleSerArgGluArgVal 277  
QY 847 AGAATCTTGAGCCCGCAGATAGAAAGAGACGTGCGCAGCAAGAGAGAAATACGATGAAGAT 906  
Db 278 ArgAlaLeuAlaProThrLysLys-----SerSer 287  
QY 907 GAATATGAATACGATGAAGAGGATAGAAGCGCTGCGAGGGGAAGCAGAGCGAGGGGAT 966  
Db 288 TrpTrpProPheGlyGlyGluSerLys----- 296  
QY 967 GGTATTGAAGACACGATCTGCACCGCNAAGTCTAAAAGAACATTTGGTAGAACAGATCC 1026  
Db 297 -----AlaGlnPheAsnIlePheSerLysArg--- 305  
QY 1027 CCGTACATCTACAACCCCTCAAGCTGTTCACTCAAAAACCTGCCAACGATCTCAACCTCTTA 1086  
Db 306 ProThrPheSerAsnGlyTyrGlyArgLeuThrGluValGlyProAspAspGluLysSer 325  
QY 1087 ATACTTAGTGGCTTGACCTAGTGTGAATATGGAATCTCTACAGGAATGCATTGTTT 1146  
Db 326 TrpLeuGlnArgLeuAsnLeuMetLeuThrPheThrAsnIleThrGlnArgSerMetSer 345  
QY 1147 GTCGCTCACTACAACCAACGACGACATCATATATCGATTGAGGCGACGGGCTCAC 1206  
Db 346 ThrIleHisTyrAsnSerHisAlaThrLysIleAlaLeuValMetAspGlyArgGlyHis 365  
QY 1207 GTGCAAGTC-----GTGGACAGCAACGCAACAGAGGTGTACACGAG----- 1248  
Db 366 LeuGlnIleSerCysProHisSerHisMetSerSerArgSerAspSerLysHisAspLysSerSer 385  
QY 1249 -----GAGCTTCAAGAGGGTCAGCTGCTTGTGGTGGCCACAG 1284  
Db 386 ProSerTyrHisArgIleSerAlaAspLeuLysProGlyMetValPheValValProPro 405  
QY 1285 AACTTCGCC---GTCGCTGMAAGTCCAGAGCGAGAACTTCGAATACGTGGCATTCAG 1341  
Db 406 GlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuLeuIleCysPheGlu 425  
QY 1342 ACAGACTCAAGGCCCGCCAGCATAGCC---AACCTCGCGGTGAAACCTCCGTCATA----- 1392  
Db 426 ValAsnValArgAspAsnLysLysPheThrPheAlaGlyLysAspAsnIleValSerSer 445









```
US-10-728-051-3 (1-1524) x US-09-323-195A-18 (1-448)
QY 4 CACAAACCGGAGAGAGCGTCCAGTTCAGCGCTCAATGCGCAGAGACCTGACAAT 63
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
46 GluGluArgGluGluAsnProTyrValPhe-----HisSerAspSerPheArgThr 62
QY 64 CGCATTGAATCAGAGGCGGTATACATTGAGACTTGAACCCCAACACAGGAGTTTCGAA 123
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
63 ArgAlaSerSerGluAlaGlyGluIleArgAlaLeu---ProAsnPheGlyGluValSer 81
QY 124 -----TGGCCGGCGTCCGCTCTCTCGTTAGTC-----CTCCGCGCAGACGCC 168
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
82 GluLeuLeuGluGlyIleArgLysPheArgValThrCysIleGluMetLysProAsnThr 101
QY 169 CTTTCGTAGGCTTTCTACTCCATGCTCCCGAGGATCTTCATCCAGCAAGGAGGGA 228
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
102 ValMetLeuProHisTyrIleAspAlaThrTrpIleLeuTyrValThrArgGlyArgGly 121
QY 229 TACTTTGGGTGTATATTCCCTGGTTCTCTAGACACTATGAGAGCCTCACACAAAGT 288
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 TyrIleAlaTyrVal-----126
QY 289 CGTCGATCTCAGTCCCAAGACACCAAGAGCTCTCAAGGAGAGAACCAAGCCAAACAG 348
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 -----HisGlnAsnGluLeuValLysArgLysLeu-----136
QY 349 CAACGAGATAGTACCAGAGGTGACCGTTTCGATGAGGTGATCTCATTCAGTTCCTC 408
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
137 -----GluGluGlyAspValPheGlyValPro 145
QY 409 ACCGGTGTGTTCTTCTGCTCTACACGACGACACTGATCTTGTGTTCTCTCTT 468
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
146 SerGlyHisThrPheTyrLeuValAsnAspHisAsnThrLeuArgIleAla---164
QY 469 ACTGACACCAACAAACAGCAACACAGTGTGATCTTCCCGAGGAGATTCATTTGGCT 528
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
164 -----164
QY 529 GGGAACACGGAGCAGAGTCTTTAAGGTACCAACAAAGCAGACAAAGCAGACGAAGA 588
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
164 -----164
QY 589 AGCTTACCATATAGCCCATACAGCCCGCAAGTACGCTTAGCAAGAGAGCGGTGATTT 648
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
165 -----SerLeuValArgProValSerThrValArgGlyGluTyrGlnProPhe 180
QY 649 AGCCCTCGAGGACAGCAGCCCGCAGAGCAGGACGAGCAAGCAAGAAACGAAGT 708
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 TyrValAlaGly-----GlyArgAsnProGln-----189
QY 709 GGAAACATCTTCAGCGCTTCACGCGGAGTCTCTGGAACAAGCCCTTCAGGTTGACGAC 768
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
190 ---ThrValTyrSerAlaPheSerAspValLeuGluAlaAlaPheAsnThrAsnVal 208
QY 769 AGACAGATAGTGCAAACCTTAAGAGCGAGCAGCCGAGAGTGAAGAGAGGAGCCATTGTG 828
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
209 GlnGlnLeuGluArgIlePheGlyGly-----HisLysSerGlyValIleIle 224
QY 829 ACAGTGGGGAGCGCTCAGATCTTGAGCCCGAGATAGAAAGAGAGCTGCCGAGGAAGA 888
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
225 -----HisAlaAsnGluGlu 229
QY 889 GAGGAATACGATGAGATGAATATGAATACGATGAAGAGGATAGAGGCTGGCAGGGA 948
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
230 -----GlnIleArgGluMetMetArgLysArgGly-----Phe 240
QY 949 AGCAGAGCGGAGGGGAATGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAAC 1008
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 SerAlaGlySerMetSerAlaProGluHis-----ProLysProPhe 254
QY 1009 ATTGGTAGAAACAGATCCCTGACATCTACACCTCAAGCTGTTCACTCAAACTGCC 1068
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
255 AsnLeuArgAsnGlnLysProAspPheGluAsn---GluAsnGlyArgPheThrIleAla 273
QY 1069 AACGATCTCAACCTTCTAATACTTAGTGGCTTGGACCTTAGCTGAATATGGAATCTC 1128
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
274 GlyProLysAsnTyrProPheLeuAspAlaLeuAspValSerValGlyLeuAlaAspLeu 293
QY 1129 TACAGGAATGCATTTGTTGTCGCTCACTACAACACCAACACAGCAGCAGCATATATCGA 1188
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
294 AsnProGlySerMetThrAlaProSerLeuAsnSerLysSerThrSerIleGlyIleVal 313
QY 1189 TTGAGGGAGCGGCTCAGTCGCAAGTCGTG-----1218
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
314 ThrAsnGlyGluGlyArgIleGluMetLacCysProHisLeuGlyGlnHisGlyTrpSer 333
QY 1219 -----GACGCAACCGCAACAGAGTGTACGACGAGGAGCTT 1254
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
334 SerProArgGluArgGlyAspGlnAspIleThrTyrGlnArgValTrp---AlaLysLeu 352
QY 1255 CAAGAGGTCACGTGCTTGTGTCACAGAACTTCGCGCTGCTGGA---AAGTCCCAAG 1311
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
353 ArgThrGlySerValTyrIleValProAlaGlyHisProIleThrGluIleAlaSerThr 372
QY 1312 ACCGAGAACTTCGATACGTGCGCATTCAGACAGACTCAAGCCCGCAGCATAGCCCAAC---1368
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
373 AsnSerArgLeuGlnIleLeuTrpPheAspLeuAsnThrArgGlyAsnGluArgGlnPhe 392
QY 1369 CTCGCGGTGAAACTCCGCTCATAGTACCTCGCGGAGAGGTGTTGCCAAATTCATAT 1428
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
393 LeuAlaGlyLysAsnValLeuAsnThrLeuGluArgGluIle-----407
QY 1429 GGCCTCCAAAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1473
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
408 -----ArgGlnLeuSerPheAsnValPro 415

RESULT 10
US-09-323-195A-17
; Sequence 17, Application US/09323195A
; Patent No. 6462257
; GENERAL INFORMATION:
; APPLICANT: Pullman, Gerald
; APPLICANT: Cairney, John
; APPLICANT: Ferreira, Ranjan
; TITLE OF INVENTION: VITILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: IPST0009
; CURRENT APPLICATION NUMBER: US/09/323,195A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 17
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Pinus taeda
US-09-323-195A-17

Alignment Scores:
Pred. No.: 2,41e-06 Length: 523
Score: 158.00 Matches: 107
Percent Similarity: 32.67% Conservative: 74
Best Local Similarity: 19.31% Mismatches: 190
Query Match: 5.87% Indels: 183
DB: 4 Gaps: 20

US-10-728-051-3 (1-1524) x US-09-323-195A-17 (1-523)
QY 10 CCGGAGGAGAACCGCTGCGCAGTTCAGCGCTCAATGCCAGAGACCTGACAAT-----63
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
37 ProGluAspHisGlyArgGlyHisGlnArgGluGluGluGluGluGluGluGluGluGluGlu 56
QY 64 -----CGCATTGAATCAGAGCGGCGGTTCACATTGACACT 96
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
57 ValPheHisSerAspArgPheArgMetArgAlaSerSerAspAlaGlyGluIleArgAla 76
```

QY 97 TGGAAACCCACACACAGGAGTTCGAA-----TGGCCCGCGTCCGCTCTCTCGCTTA 150  
Db 77 Leu---ProAsnPheGlyGluAlaSerGluLeuLeuGluGlyIleSerIleValArgVal 95  
QY 151 GTC-----CTCCGCCCAACCGCTTCGTAGGCTTTCTACTCCAATGCTCCCCAG 201  
Db 96 ThrCysIleGluMetArgProAsnThrValMetLeuProHisIleValLeuAspAlaThrTrp 115  
QY 202 GAGATCTTCATCCAGCAAGAGGGGATCTTTGGGTGGATATTCCTGGTCTGCTCCTAGA 261  
Db 116 IleLeuTyrValThrGlyGlyArgGlyTyrIleAlaTyrVal----- 129  
QY 262 CACTATGAAGAGCTCACACACAGAGTTCGTCTCAGTCCCAAGACACCAAGACGCT 321  
Db 130 -----HisGlnAsnGluLeuValIleValIleValIleValIleValIleVal 138  
QY 322 CTCAAGGAGAGACCAAGACCAACAGACGAGATAGTCAACAGAGGTGCACCGCTTC 381  
Db 139 Leu----- 139  
QY 382 GATGAGGTGATCTCATGTCAGTTCCACCGGTGTTGCTTTCTGGCTCTCAACAGCAC 441  
Db 140 GluGluGlyAspValPheGlyValProSerGlyHisThrPheTyrLeuValAsnAsp 159  
QY 442 GACACTGATGTTGCTGTTCTCTTACTGACACCAACCAACAGACACACGCTTGAT 501  
Db 160 AspHisAsnSerLeuArgIleThr----- 167  
QY 502 CAGTTCCTCCAGGAGATTCAATTTGGCTGGGAACACGAGCAAGATTCTTAAGTACCAG 561  
Db 168 -----SerLeuLeuValThrVal 173  
QY 562 CACAAAGCAGACAAAGCAGACGAGAGAGTTACCATATAGCCCATACAGCCCGCAAGT 621  
Db 174 SerThrMetArgGlyGlu-----TyrGluProTyr----- 183  
QY 622 CAGCCTAGACAAAGAGCGTGAATTTAGCCCTCGAGCAGACGACCGCCAGACGA 681  
Db 184 -----TyrValAla 186  
QY 682 GCAGGCAAGAAAGAAACGAGGTGGAAACATCTTCAGCGCTTCACGCGGAGTTC 741  
Db 187 GlyIleArgAsnProGlu-----ThrValTyrSerAlaPheSerAspVal 202  
QY 742 CTGGAACAGCCTTCAGTTGACGACGACAGATAGTCAAAACCTTAAGAGCGAGACC 801  
Db 203 LeuGluAlaAlaPheAsnThrAsn----- 210  
QY 802 GAGTGAAGAGGAGGCGCATGTGACAGTGGGGGCGCTCAGAACTTTGAGCCCA 861  
Db 211 -----ValIleGluAlaArgThrHisPheProVal----- 220  
QY 862 GATGAAGAGAGCGTCCGCGAGAGAGAGATACGATGAAGATGAATGAATACGAT 921  
Db 221 -----HisIleGluArgGluSerTyrSerMetAla 230  
QY 922 GAAGAGATGAAGCGCTGGCAGGGAAGCAGAGCGAGGGGAATGGTATTTGAAGAGCG 981  
Db 231 AsnGluGluGlnIleArg---GluMetLeuArgIleValPheSerAlaGluSerMet 249  
QY 982 ATCTGCACCCCAAGTGTAAAGAACATTTGGTAGAACAGATCCCTGATCATCTACAAC 1041  
Db 250 SerAlaSerGluHisProIleValLeuAsnProGlyThrMetThrAlaProSerHisAsn 269  
QY 1042 CCTCAAGCTGTTCACTCAAACTGCCAAGCATCTCAACTCTTAATACCTTAGTGCTT 1101  
Db 270 AsnAsnGlyArgPheThrArgAlaGlyProAsnGluAsnProLeu---LeuAspAlaVal 288  
QY 1102 GGACCTAGTCTGAATATGAAATCTCTACAGGAATGCATTTGTTGCTCTCACTACAAC 1161  
Db 289 AspValThrAlaGlyPheGlyValLeuAsnProGlyThrMetThrAlaProSerHisAsn 308  
QY 1162 ACCAACCCACACGATCATATATCGATTGAGGGGACGGCTCACGTGCAAGTCGTG--- 1218

Db 309 ThrLysAlaThrSerIleAlaIleValThrGlnGlyGluGlyArgIleGluMetAlaCys 328  
QY 1219 -----GACAGCAAC 1227  
Db 329 ProHisLeuGlyGlnHisGlyTyrSerSerArgGluLysGlyAspGlnGluIleAsn 348  
QY 1228 GGCACACAGAGTGTACACGAGGAGCTTCAAGAGGGTCAAGTGTGTTGGTGGCCACAGAAC 1287  
Db 349 TyrGlnArgVal---ArgAlaArgLeuArgThrGlyThrValTyrValValProAlaGly 367  
QY 1288 TTCGCCGTGCTGCTGAAAGTCCAGAGCGAG---AACTTCGAATACGTGGCATTCAGACACA 1344  
Db 368 HisProIleThrGluIleAlaCysThrGluGlyHisLeuGluIleLeuTrpPheAspIle 387  
QY 1345 GACTCAAGGCCACGATAGCCAAAC---CTGCCGCGTGAAACCTCCGTCATAGATAACCTG 1401  
Db 388 AsnThrSerGlyAsnGluArgGlnPheLeuAlaGlyLysTyrAsnValLeuGlnThrLeu 407  
QY 1402 CCGGAGAGGTGTTGCTCAATTCATATGCTCCAAAGG-----GACCAG 1446  
Db 408 GluLysGluValArgGlnIleSerPheAsnIleProArgGlyGluGluLeuAspGluVal 427  
QY 1447 GCAAGCGAGCTTAAGAAC-----AACAAAC 1470  
Db 428 LeuArgArgGlnLysAspGlnValIleLeuArgGlyProGlnMetGlnArgArgAspGlu 447  
QY 1471 CCCTTCAAGTTCTTCGTTCCACCGCTCTCAGCAGCTCTCCGA 1510  
Db 448 -ProArgSerSerIleHisAlaIleIleAlaAlaArg 460  
RESULT 11  
US-09-106-872A-17  
; Sequence 17, Application US/09106872A  
; Patent No. 6486311  
; GENERAL INFORMATION:  
; APPLICANT: Burks Jr., A. Wesley  
; APPLICANT: Stanley, J. Steven  
; APPLICANT: Cockrell, Gael  
; APPLICANT: King, Nina E.  
; APPLICANT: Sampson, Hugh A.  
; APPLICANT: Helm, Ricki M.  
; APPLICANT: Bannan, Gary A.  
; TITLE OF INVENTION: Peanut Allergens and Methods  
; FILE REFERENCE: HS 103 CIP  
; CURRENT APPLICATION NUMBER: US/09/106,872A  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: PCT/US96/15222  
; PRIOR FILING DATE: 1996-09-23  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 335  
; TYPE: PRT  
; ORGANISM: Arachis hypogaea  
US-09-106-872A-17  
Alignment Scores:  
Pred. No.: 2,14e-06 Length: 335  
Score: 157.50 Matches: 55  
Percent Similarity: 42.04% Conservative: 48  
Best Local Similarity: 22.45% Mismatches: 81  
Query Match: 5.85% Indels: 61  
Gaps: 9  
US-10-728-051-3 (1-1524) x US-09-106-872A-17 (1-335)  
QY 31 TTCAGGCGCTCAATCGCGAGACCTGACAATCGCATTCAGAGGGCGGTTCATT 90  
Db 122 PheGlyLysLeuPheGluValIleValProAspLysLys----- 133  
QY 91 GAGACTTGGAAACCCCAACACAGAGGAGTTCAATGCGCGCGCTCGCCCTCTCTCGCTTA 150  
Db 91 GAGACTTGGAAACCCCAACACAGAGGAGTTCAATGCGCGCGCTCGCCCTCTCTCGCTTA 150



```
Db 256 HisGluAsnLysGluSerTyrAsnValValProGlyValValValArgValProAlaGly 275
Qy 844 -----CTCAGATCTTGAGCCAGATAGAAAG----- 870
Db 276 SerThrValTyrLeuAlaAsnGlnAspAsnLysGluLysLeuIleAlaValLeuHis 295
Qy 871 AGACGTGCCACGACGAAGAGGATACGATGAA----- 903
Db 296 ArgProValAsnAsnProArgGlnPheGluGluPhePheProAlaGlySerGlnArgPro 315
Qy 903 ----- 903
Db 316 GlnSerTyrLeuArgAlaPheSerArgGluIleLeuGluProAlaPheAsnThrArgSer 335
Qy 904 -----GATGAATATGATACGATGAAGAGATAGAGCGGTGCACGGGGAAGCAGA 954
Db 336 GluGlnLeuAspGluLeuPheGlyGlyArgGlnSerHisArgGlnGlnGlnGlyGlnGly 355
Qy 955 GGCAGGGGGGATGGTATTGAGAGCAGCATCTGCACCCAGATGCTAAAGAGACATTGGT 1014
Db 356 MetPheArgLysAlaSerGlnGlnGlnIleArgAlaLeuSerGlnGlnAlaThrSerPro 375
Qy 1015 AGAAACAGATCCCTCTGAC-----ATCTACAAC 1041
Db 376 ArgGluLysSerGlyGluArgPheAlaPheAsnLeuTyrArgThrProArgTyrSer 395
Qy 1042 CTTCAAGCTGGT-----TCACCTCAAACTGCCACGATCTC 1077
Db 396 AsnGlnAsnGlyArgPheTyrGluAlaCysProArgGluPheArgGlnLeuSerAspIle 415
Qy 1078 AACCTTCTAATCTAGTGGCTGGACCTAGCTGCTCAATATGGAATCTCTACAGGAAT 1137
Db 416 AsnValThrVal-----SerAlaLeu-----GlnLeuAsnGlnGly 427
Qy 1138 GCATTGTTTCGCTCACTACAACACCAACGACACACAGCATATATCGATTGAGGGGA 1197
Db 428 SerIlePheValProHisTyrAsnSerLysAlaThrPheValValLeuValAsnGluGly 447
Qy 1198 CGGGCTCAGTCGAAGTCGTGGAC----- 1221
Db 448 AsnGlyTyrValGluMetValSerProHisLeuProArgGlnSerSerPheGluGlu 467
Qy 1222 -----AGCAACGGCAACAGGTG 1239
Db 468 GluGlnGlnGlnGlnGlnGlnGlnGluArgArgSerGlyGlnTyrArgLys 487
Qy 1240 TAGCAGCAGAGGCTTCAAGAGGGTCACGTGTTGTGGTCCACAGAACTTCGCGCTCGCT 1299
Db 488 IleArgSerGlnLeuSerArgGlyAspIlePheValValProAlaAsnPheProValThr 507
Qy 1300 ---GGAAAGTCCCAGACGAGAACTTCGAATACGTGGCATTC----- 1338
Db 508 PheValAlaSerGlnAsnGlnAsnLeuArgMetThrGlyPheGlyLeuTyrAsnGlnAsn 527
Qy 1339 ---NAGACACTCAAGCCCGCCAGTATAGCACTCGCGGTGNAAACTCCGTATAGAT 1395
Db 528 IleAsnProAspHisAsnGlnArgIlePhe---ValAlaGlyLysIleAsnHisValArg 546
Qy 1396 AACCTGCGGAGGAGGTGGTTGCAAAATTCATATGCGCTCCAAAGGGAGCAGCAAGCCAG 1455
Db 547 GlnTrpAspSerGlnAlaLysGluLeuAlaPheGlyValSerSerArgLeuValAspGlu 566
Qy 1456 CTTAAGAACCAACCCCTCAAG---TTCTTGCTTCCACCGTCTCAGCAGTCTCCGAGG 1512
Db 567 IlePheAsnAsnProGlnGluSerTyrPheVal-----SerArgGlnArgGlnArg 584
Qy 1513 GCT 1515
Db 585 Ala 585
```

RESULT 13

US-07-955-905A-2

; Sequence 2, Application US/07955905A

```
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (SPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-955-905A-2
```

## Alignment Scores:

```
Pred. No.: 3 91e-06 Length: 566
Score: 156.00 Matches: 106
Percent Similarity: 34.82% Conservative: 74
Best Local Similarity: 20.50% Mismatches: 205
Query Match: 5.80% Indels: 132
DB: 1 Gaps: 22
```

US-10-728-051-3 (1-1524) x US-07-955-905A-2 (1-566)

```
Qy 253 TGTCTAGACACTATGAAGAGCCTCACACAAGAGTCGTGCTCTCAGTCCCAAGACCA 312
Db 66 CysGluArgGluTyrLysGluGlnArgGln---GlnGluGluGluLeuGlnArgGln 84
Qy 313 CCAAGAGCTCTCAAGAGAGACCAAGACCAAGCCACAGCAA----- 351
Db 85 TyrGlnGlnCysGlnGlyArgCysGlnGlnGlnGlnGlnGlnGlnGln 104
Qy 352 -----CGAGATAGTCACCAAGAGGTG 372
Db 105 CysGlnArgLysCysTrpGluGlnTyrLysGluGlnGluArgGlyGluHisGluAsnTyr 124
Qy 373 CACCGTTTC-----GATGAGCGT-----GATCTCATT 399
Db 125 HisAsnHisLysLysAsnArgSerGluGluGluGluGlnGlnArgAsnAsnProTyr 144
Qy 400 GCAGTCCCAACCGGTGTTGCTTTCTGGCTC---TACAAACGACACGACACTGATGTTGT 456
Db 145 TyrPheProLysArgArgSerPheGlnThrArgPheArgAspGluGluGlyAsnPheLys 164
Qy 457 GCTGTTTCTCTTACTGACCAACAAC-----AACGACAAACAGCTT 498
Db 165 IleLeuGlnArgPheAlaGluAsnSerProPheLeuLysGlyIleAsnAspTyrArgLeu 184
Qy 499 GATCAGTTC-----CCCAGGAGATTCAATTTGGTGGGAACACGACGAGCAAGAG--- 546
Db 185 AlaMetPheGluAlaAsnProAsnThrPheIleLeuProHisCysAspAlaGluAla 204
Qy 547 -----TTCTTAAGGTACCAACCAACAAGC 570
Db 205 IleTyrPheValThrAsnGlyLysGlyThrIleThrPheValThrHisGluAsnLysGlu 224
Qy 571 AGACAAGACAGACGAAGA-----AGCTTACCATAT---AGCCCATACAGCCCG 615
Db 225 SerTyrAsnValGlnArgGlyThrValValSerValProAlaGlySerThrValTyrVal 244
Qy 616 CAAAGTCAGCCTACAGAAAGAGCGTGAATTTAGC-----CCTCGAGGACAG 663
Db 245 ValSerGlnAspAsnGlnGluLysLeuThrIleAlaValLeuAlaLeuProValAsnSer 264
```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Theobroma cacao
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..566
; OTHER INFORMATION: /note= "67 kD Precursor Protein"
;
US-07-955-905A-22

```

```

;
; Alignment Scores:
; Pred. No.: 3,91e-06 Length: 566
; Score: 156.00 Matches: 106
; Percent Similarity: 34.82% Conservative: 74
; Best Local Similarity: 20.50% Mismatches: 205
; Query Match: 5.80% Indels: 132
; DB: 1 Gaps: 22
;
US-10-728-051-3 (1-1524) x US-07-955-905A-22 (1-566)

```

```

;
; 253 TGTCCTAGACACTATGAAGAGCCTCACACACAGGTGCTGATCTCAGTCCCAAGACCA 312
; 66 CysGluArgGluTyrLysGluGlnGlnArgGln---GlnGluGluGluGlnArgGln 84
; 313 CCACAGCTCTCCAGAGGAGAGACCAAGCCCAACAGCAA----- 351
; 85 TyrGlnGlnCysGlnGlyArgCysGlnGluGlnGlnGlnGlnGlnGlnGln 104
; 352 -----CGAGATAGTCACCAAGAGTG 372
; 105 CysGlnArgLysCysTyrGlnTyrLysGluGlnGluArgGlyGluHisGluAsnTyr 124
; 373 CACCGTTTC-----GATGAGGT-----GATCTCAT 399
; 125 HisAsnHisLysLysAsnArgSerGluGluGluGluGlnGlnArgAsnAsnProTyr 144
; 400 GCAGTTCCACCGGTGTTGCTTTCTGGCTC---TACAACGACCAACGACGACTGTTGTT 456
; 145 TyrPheProLysArgSerPheGlnThrArgPheArgAspGluGluGlyAsnPheLys 164
; 457 GCTGTTTCTTACTGACACCAACAC-----AACGACACACAGCTT 498
; 165 IleuGlnArgPheAlaGluAsnSerProProLeuLysGlyIleAsnAspTyrArgLeu 184
; 499 GATCAGTTC-----CCGAGGATTCATTTGGTGGGAACACGAGCAAGAG--- 546
; 185 AlaMetPheGluAlaAsnProAsnThrPheIleLeuProHisCysAspAlaGluAla 204
; 547 -----TTCTTAAGTTACCAACCAAGAGC 570
; 205 IleTyrPheValThrAsnGlyLysGlyThrIleThrPheValThrHisGluAsnLysGlu 224
; 571 AGACAAGACGACGAGA-----AGCTTACCATAT---AGCCCATACACCCCG 615
; 225 SerTyrAsnValGlnArgGlyThrValSerValProAlaGlySerThrValTyrVal 244
; 616 CAAAGTCAGGCTAGACCAAGAGAGCGGTGAATTTAGC-----CCTCAGAGACAG 663
; 245 ValSerGlnAspAsnGlnGluLysLeuThrIleAlaValLeuAlaLeuProValAsnSer 264

```

```

;
; RESULT 14
; US-07-955-905A-22
; Sequence 22, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28

```

```
QY 664 CACAGCCGACGAGACGAGCAGGACCAAGAAAGAAACGAGGTGGAACATCTTCAGC 723
Db 265 ProGlyLysTyrGluLeuPheProAlaGlyAsnAsnLysProGluSerTyrGly 284
QY 724 GGCTTCACGCGGAGTTCCTGGAACAGCCTTCAGGTTGACGAC-----AGACAG 774
Db 285 AlaPheSerTyrGluValLeuGluThrValPheAsnThrGlnArgGluLysLeuGluGlu 304
QY 775 ATAGTGCAAACCTTAAGCGGAGCCGAGAGTGAAGAGAGGAGGAGCCATTGTGACAGTG 834
Db 305 IleLeuGluGluGlnArgGlyGlnLysArgGlnGlnGlyGln----- 318
QY 835 AGGGAGCGCTCGAATCTTGAGCCGACATAGAAGAGAGCTGCCGACGAGAGAGGAA 894
Db 319 GlnGlyMetPheArgLysAlaLysProGluGlnIleArgAlaIleSerGlnAla--- 337
QY 895 TACGATGAAGATGAATATGAATACGATGAAGAGAGGATAGAAGGCGTGCAGGGAAGCAGA 954
Db 338 -----ThrSerProArgHisArgGly---GlyGluArg 347
QY 955 GGCAGGGGGGATGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAAGACATTGGT 1014
Db 348 LeuAlaIleAsnLeuLeu----- 353
QY 1015 AGAACAGATCCCTGACATCTACAACCTCAAGCTGGTTCACTCAAAACTGCCAACGAT 1074
Db 354 ---SerGlnSerPro---ValTyrSerAsnGlnArgPhePheGluAlaCysPro 371
QY 1075 CTCAACTCTTAATCTAGTGGCTGGAGCTAGTCTGAATATGGAATCTCTACAGG 1134
Db 372 GluAspPheSerGlnPheGlnAsnMetAspValAlaValSerAlaPheLysLeuAsnGln 391
QY 1135 AATGCATTGTTGCTCTACTACAACCAAGCAGACATCATATATCGATTGAGG 1194
Db 392 GlyAlaIlePheValProHisTyrAsnSerLysAlaThrPheValPheValThrAsp 411
QY 1195 GACGGGCTCAGTGCAGTCTG-----GACAGCAAGCGCAAC--- 1233
Db 412 GlyTyrGlyTyrAlaGlnMetAlaCysProHisLeuSerArgGlnSerGlnGlySerGln 431
QY 1234 -----AGAGTGTACGACGAGGAGCTTCAAGAG----- 1260
Db 432 SerGlyArgGlnAspArgGluGlnGluGluSerGluGluThrPheGlyGlu 451
QY 1261 -----GGTCACGTGTTGTGTGTCACAGAACTTC 1290
Db 452 PheGlnGlnValLysAlaProLeuSerProGlyAspValPheValAlaProAlaGlyHis 471
QY 1291 GCGGTC-----GCTGGAAGTCCGACGAGGAACTTCGATACGTTGCAAG 1341
Db 472 AlaValThrPhePheAlaSerLysAspGlnProLeuAsnAlaValAlaPheGlyLeuAsn 491
QY 1342 ACAGACTCAAGGCGCCACATAGCAACCTCGCGGTGAAATCCGTCATAGATAAAGCTG 1401
Db 492 AlaGlnAsnGlnArgIlePhe---LeuAlaGlyLysLysAsnLeuValArgGlnMet 510
QY 1402 CCGGAGGAGTGGTGTGCAAAATTCATATGGCTCCAAAGGAGGAGGAGGCAAGCGCTTAAG 1461
Db 511 AspSerGluAlaLysGluLeuSerPheGlyValProSerLysLeuValAspAsnIlePhe 530
QY 1462 AACACAGACCCCTCAAGTCTTCGTTCCACCGTCTCAGCAGTCTCCGAGG 1512
Db 531 AsnAsnProAspGluSerTyrPheMetSerPheSerGlnGlnArgGlnArg 547
RESULT 15
US-09-216-393B-81
; Sequence 81, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milhaussen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
```

```
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 81
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
; US-09-216-393B-81

Alignment Scores:
Pred. No.: 4,07e-06 Length: 611
Score: 150.00 Matches: 92
Percent Similarity: 35.11% Conservative: 73
Best Local Similarity: 19.57% Mismatches: 169
Query Match: 5.80% Indels: 137
DB: 4 Gaps: 14

US-10-728-051-3 (1-1524) x US-09-216-393B-81 (1-611)
QY 29 AGTTCACGCGCCTCAATGCGCAGACCTGACAATCGCATTGAATCAGAGGGCGGTACA 88
Db 216 SerSerProIleSerValThrAsnAspValTyrAlaLeuPheAspSerSerAlaSerPro 235
QY 89 TTGAGACTTGGAAACCCCAACACACAGGAGTTGGAATGCGCGCGGCTCTCTCGCT 148
Db 236 LeuHisAlaGlyGluLeuSerSerLeuProGlyAlaValSerAlaSerGluArgLeu-Le 255
QY 149 TACTCTCCG-----CCGCAACGCCCTTCGTAGGCCCTTCTACTCCAATCCTC 196
Db 255 uThrAlaProAlaGluIleGlyProSerAlaSerSerAlaCysLeuSerValSerCysG 275
QY 197 CCACGAGATCTTCATCCAGCAAGGAGGATCTTTGGGTTGATATTCCTCGGTTGTC 256
Db 275 yProGlyGluMetSerProThrAlaAsp-----ThrTh 286
QY 257 CTAGACAC-----TATGAAGAGCTTCACACAAAGGTC 289
Db 286 r-ArgHisAspAlaGluGluArgAlaGluGluLysGluArgGlu 306
QY 290 GTCGATCTCAGTCCCAAGACCAACAGAGCTTCCAAAGGAGAGACCAAGCAACAGC 349
Db 306 rGlnGluGluGluGluArgGluArgValGluGluGluGluGluGluArgG 326
QY 350 AACGAGATAGTACCAAGAGGTGCACCGCTTTCGATGAGGTGATCTCATTCGATTCCTCA 409
Db 326 lndGluGluGluArgGluArgValGlu----- 338
QY 410 CCGGTGTTGTTTCTGCTCTACACGACCAACGACTGATGTTGTTGTTTCTCTTA 469
Db 338 ----- 338
QY 470 CTGACACCAACAACAGCAACAGCTTTCAGTTCCCGCAGGAGTTCATTTGGCTG 529
Db 339 -----GluLysAlaArgGlnArgGluGluArgGluArgGluArgG 352
QY 530 GGAACACGAGCAAGAGTCTTAAGGTACCAGCAACAAGCAGACAGCAAGCAGCAAGAA 589
Db 353 --ArgValGluGluGluLysAlaArgGlnArgGluGluGluGluArgGluArgG- 371
QY 590 GCTTACATATAGCCCATACAGCCCGCAAGTCAAGCTAGACAAAGAGCGGTGAATTTA 649
Db 372 -----ValGluGluGluLysAlaArgGlnArgGluGluGluGluArgG 387
QY 650 GCCTCTGAGACACAGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 709
Db 387 luArgArgValGluGluGluLysAlaArgGlnArgGluGluGluGluGluArgGlu 407
QY 710 GAAACATCTTCAGCGGCTTCACCGCGAGTTCCTGGAACAGGCTTCACGTTGACGACA 769
Db 407 rgArgArgValGluGluGluLysAlaArgGlnArgGluGluGluGluGluArgGluGly 427
```



Search completed: April 8, 2005, 05:34:42  
Job time : 72.194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 8, 2005, 05:25:59 ; Search time 174.941 Seconds  
(without alignments)  
5784.401 Million cell updates/sec

Title: US-10-728-051-3  
Perfect score: 2691  
Sequence: 1 cgcagcaaccgaggagaa.....ctccgaggctgtggttaa 1524

Scoring table: BLOSUM62  
Ggapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 2636020

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-O=/cgn2 1/USPTO spoel p/US10728051/runat 07042005 125048 14128/app query.fasta\_1.4757  
-DB=Published Applications AA -QPMT=fastan -SUFFIX=rapb -WINMATCH=0.1  
-LOOPCUT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10728051 @CGN 1 1 663 @runat 07042005 125048 14128  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: /cgn2 6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp:\*  
2: /cgn2 6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pcp:\*  
3: /cgn2 6/ptodata/1/pubpaa/US06\_NEW\_PUB.pcp:\*  
4: /cgn2 6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp:\*  
5: /cgn2 6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp:\*  
6: /cgn2 6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp:\*  
7: /cgn2 6/ptodata/1/pubpaa/US08\_NEW\_PUB.pcp:\*  
8: /cgn2 6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp:\*  
9: /cgn2 6/ptodata/1/pubpaa/US09A\_PUBCOMB.pcp:\*  
10: /cgn2 6/ptodata/1/pubpaa/US09B\_PUBCOMB.pcp:\*  
11: /cgn2 6/ptodata/1/pubpaa/US09C\_PUBCOMB.pcp:\*  
12: /cgn2 6/ptodata/1/pubpaa/US09\_NEW\_PUB.pcp:\*  
13: /cgn2 6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp:\*  
14: /cgn2 6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp:\*  
15: /cgn2 6/ptodata/1/pubpaa/US10C\_PUBCOMB.pcp:\*  
16: /cgn2 6/ptodata/1/pubpaa/US10D\_PUBCOMB.pcp:\*  
17: /cgn2 6/ptodata/1/pubpaa/US10\_NEW\_PUB.pcp:\*  
18: /cgn2 6/ptodata/1/pubpaa/US11\_NEW\_PUB.pcp:\*  
19: /cgn2 6/ptodata/1/pubpaa/US60\_NEW\_PUB.pcp:\*  
20: /cgn2 6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	2673	99.3	507	15	US-10-245-871-96	Sequence 96, Appl
2	2673	99.3	507	15	US-10-253-286-96	Sequence 96, Appl
3	2673	99.3	510	14	US-10-228-806-6	Sequence 6, Appl
4	2673	99.3	510	15	US-10-100-303A-90	Sequence 90, Appl
5	2638	98.0	526	9	US-09-731-221-79	Sequence 79, Appl
6	2621	97.4	530	17	US-10-899-551-6	Sequence 6, Appl
7	1665	61.9	351	17	US-10-899-551-58	Sequence 58, Appl
8	1502	55.8	481	10	US-09-759-967-21	Sequence 21, Appl
9	1502	55.8	481	15	US-10-424-599-171702	Sequence 171702,
10	1502	55.8	488	15	US-10-425-114-43865	Sequence 43865, A
11	1502	55.8	491	15	US-10-425-114-44047	Sequence 44047, A
12	1502	55.8	492	15	US-10-425-114-43836	Sequence 43836, A
13	1502	55.8	492	15	US-10-425-114-46394	Sequence 46294, A
14	1499.5	55.7	466	15	US-10-425-114-71959	Sequence 71959, A
15	1499.5	55.7	466	15	US-10-425-114-71973	Sequence 71973, A
16	1499.5	55.7	485	10	US-09-759-967-20	Sequence 20, Appl
17	1499.5	55.7	488	15	US-10-425-114-45645	Sequence 45645, A
18	1499.5	55.7	489	15	US-10-425-114-43983	Sequence 43983, A
19	1499.5	55.7	489	15	US-10-425-114-44676	Sequence 44676, A
20	1499.5	55.7	489	15	US-10-425-114-45612	Sequence 45612, A
21	1499.5	55.7	489	15	US-10-425-114-50117	Sequence 50117, A
22	1499.5	55.7	489	15	US-10-425-114-50207	Sequence 50207, A
23	1499.5	55.7	489	15	US-10-425-114-51412	Sequence 51412, A
24	1499.5	55.7	489	15	US-10-425-114-58429	Sequence 58429, A
25	1499.5	55.7	489	15	US-10-425-114-58429	Sequence 71902, A
26	1499.5	55.7	489	15	US-10-425-114-71909	Sequence 71909, A
27	1499.5	55.7	489	15	US-10-425-114-71957	Sequence 71957, A
28	1499.5	55.7	489	15	US-10-425-114-71958	Sequence 71958, A
29	1499.5	55.7	489	15	US-10-425-114-71967	Sequence 71967, A
30	1499.5	55.7	489	15	US-10-425-114-71995	Sequence 71995, A
31	1499.5	55.7	489	15	US-10-425-114-72287	Sequence 72287, A
32	1499.5	55.7	489	15	US-10-425-114-72288	Sequence 72288, A
33	1498	55.7	489	15	US-10-425-114-45639	Sequence 45639, A
34	1496	55.6	495	10	US-09-759-967-2	Sequence 2, Appl
35	1496	55.6	495	15	US-10-424-599-171705	Sequence 171705,
36	1496	55.6	501	15	US-10-425-114-45626	Sequence 45626, A
37	1496	55.6	501	15	US-10-425-114-43904	Sequence 43904, A
38	1496	55.6	505	15	US-10-425-114-71930	Sequence 71930, A
39	1496	55.6	505	15	US-10-425-114-71962	Sequence 71962, A
40	1496	55.6	506	15	US-10-425-114-43901	Sequence 43901, A
41	1496	55.6	506	15	US-10-425-114-45624	Sequence 45624, A
42	1496	55.6	506	15	US-10-425-114-45680	Sequence 45680, A
43	1496	55.6	507	15	US-10-425-114-43929	Sequence 43929, A
44	1496	55.6	507	15	US-10-425-114-45765	Sequence 45765, A
45	1496	55.6	508	15	US-10-425-114-45325	Sequence 45325, A

ALIGNMENTS

RESULT 1  
US-10-245-871-96  
; Sequence 96, Application US/10245871  
; Publication No. US20030235594A1  
; GENERAL INFORMATION:  
; APPLICANT: HUMPHREYS, ROBERT  
; APPLICANT: XU, MINZHEN  
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES  
; FILE REFERENCE: REH-2013  
; CURRENT APPLICATION NUMBER: US/10/245, 871  
; CURRENT FILING DATE: 2003-01-09  
; PRIOR APPLICATION NUMBER: 10/197,000  
; PRIOR FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: 09/396,813  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 905  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 96  
; LENGTH: 507  
; TYPE: PRT  
; ORGANISM: Arachis hypogaea  
US-10-245-871-96

Alignment Scores: 1.55e-221 Length: 507  
Pred. No.: 2673.00 Matches: 507  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 99.33% Gaps: 0  
DB: 15

US-10-728-051-3 (1-1524) x US-10-245-871-96 (1-507)

QY 1 CGGCAGCAACCGGAGGAGAACGGTCCAGTTCAGCGCTCAATGCGCAGAGACTGAC 60  
DB 1 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 20  
QY 61 AATCGCATTAATCAGAGGCGGTTTACATTGAGACTTGGAAACCCCAACACAGAGATTTC 120  
DB 21 AsnArgilleGluSerGluGlyGlyTyrilleGluThrTrpAsnProAsnAsnGlnGluPhe 40  
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180  
DB 41 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 60  
QY 181 TTCTACTCCATGCTCCCAAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGGTTG 240  
DB 61 PheTyrSerAsnAlaProGlnGluillePheilleGlnGlnGlyArgGlyTyrPheGlyLeu 80  
QY 241 ATATTCCCTGGTGTCTAGACACTATGAGAGCCTCACACAGGTGTCGATCTCAG 300  
DB 81 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 100  
QY 301 TCCCAAGACACCAAGACCTCCCAAGGAGAGACCCAAAGCCAAACAGCAACAGATAGT 360  
DB 101 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 120  
QY 361 CACCAAGAGTGCAACCTTTTCGATAGGGTGAATCTATTGAGTTCACCGGTGTTGCT 420  
DB 121 HisGlnLysValHisArgPheAspGluGlyAspLeuileAlaValProThrGlyValAla 140  
QY 421 TTCTGGCTTACACGACACCGACACTGATGTTGCTGTTCTTCTTACTGACACCAAC 480  
DB 141 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 160  
QY 481 AACACAGCAACACAGCTTGATTCAGTCCCGCAGGAGATTCATTTGGCTGGGAACAGGAG 540  
DB 161 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180  
QY 541 CAAGAGTCTTAAGGTACCGCAACAAAGCGACAAAGCAGCAAGCAAGCTTACCATAT 600  
DB 181 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 200  
QY 601 AGCCCATACCGCGCAAGTCAGCTAGACCAAGAGCGTGAATTTAGCCCTCCAGGA 660  
DB 201 SerProTyrSerProGlnSerGlnProArgGlnGlnGluArgGluPheSerProArgGly 220  
QY 661 CAGCACGCCCGCAGAGACGAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAATCTTC 720  
DB 221 GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluAsnGluGlyGlyAsnillePhe 240  
QY 721 AGCGCTTTCAGCGGAGTTCCTGGAAACAGCCCTTCAGGTTGACGACAGACAGATAGTG 780  
DB 241 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnleVal 260  
QY 781 CAACAACCTTAAGGGGACCGCAGAGTCAAGAGAGGAGGAGCATTTGTGACAGTCCAGGGA 840  
DB 261 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlalleValThrValArgGly 280  
QY 841 GGCCTCAGAAATCTTGAGCCGACATAGAAGACAGTCCCGACGAAAGAGAGAAATACAT 900  
DB 281 GlyLeuArgilleLeuSerProAspArgLysArgArgAlaAspGluGluGluGluTyrAsp 300  
QY 901 GAAGATGATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960

301 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 320  
961 GGGAAATGTTATTAAGAGAGAGATCTGACCGCGCAGTCTAAAAAGAACATTTGTTAGAAAC 1020  
321 GlyAsnGlyilleGluGluThrilleCysThrAlaSerAlaLysLysAsnilleGlyArgAsn 340  
1021 AGATCCCTGACATCTACACCTCAAGCTGTTGTTCACTCAAACTGCCAACGATCTCAAC 1080  
341 ArgSerProAspilleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 360  
1081 CTTCTAATACTAGTTCGCTTGGACCTAGTCTCAATATGAAATCTCTACAGGAATGCA 1140  
361 LeuLeuilleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 380  
1141 TTGTTTTCGCTCACTCAACACCAACGCAACACAGATCATATATTCGATTGAGGGACGG 1200  
381 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerilleTyrArgLeuArgGlyArg 400  
1201 GCTCAGTGCAGAGTCTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAG 1260  
401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 420  
1261 GGTCAAGTCTGTTGTCGCCACAGAACTTCGCGCTCGCTGGAAAGTCCACAGAGGAGAAC 1320  
421 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 440  
1321 TTGCAATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCAACCTCGCGGTGAA 1380  
441 PheGluTyrValAlaPheLysThrAspSerArgProSerilleAlaAsnLeuAlaGlyGlu 460  
1381 AACTCGCTCATAGATAACTGCGGAGGAGGTGTTGCATATTCATATGCGCTCCCAAGG 1440  
461 AsnSerValilleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 480  
1441 GAGCAGGCAAGCAGCTTAAGAAACAAACCCCTTCAAGTTCTTTCGTTCCACCGTCTCAG 1500  
481 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProSerGln 500  
1501 CAGTCTCCGAGGCTGTGGCT 1521  
501 GlnSerProArgAlaValAla 507

RESULT 2  
US-10-253-286-96  
; Sequence 96, Application US/10253286  
; Publication No. US20040058881A1  
; GENERAL INFORMATION:  
; APPLICANT: HUMPHREYS, ROBERT  
; APPLICANT: XU, MINZHEN  
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES  
; FILE REFERENCE: REH-2015  
; CURRENT APPLICATION NUMBER: US/10/253,286  
; CURRENT FILING DATE: 2003-01-13  
; PRIOR APPLICATION NUMBER: 10/197,000  
; PRIOR FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: 09/396,813  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 905  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 96  
; LENGTH: 507  
; TYPE: PRT  
; ORGANISM: Arachis hypogaea  
US-10-253-286-96

Alignment Scores: 1.55e-221 Length: 507  
Pred. No.: 2673.00 Matches: 507  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 99.33% Gaps: 0  
DB: 15

US-10-728-051-3 (1-1524) x US-10-253-286-96 (1-507)

QY 1 CGGAGAACCGGAGGAGAACGGCTGCCAGTTCAGCGCCTCAATGCGCAGACCTTGAC 60  
Db 1 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 20

QY 61 AATCGCATTCGAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACCGAGGATTC 120  
Db 21 AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe 40

QY 121 GAATGCGCGCGGCTGCGCCTCTCTCGCTTAGTCTCCGCGCAGAACCGCCTTCGTAGGCGCT 180  
Db 41 GluCysalaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 60

QY 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGAGGATCTTTGGGTTG 240  
Db 61 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 80

QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCTCACACAAAGGTGCTCGATCTCAG 300  
Db 81 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgGserGln 100

QY 301 TCCCAAGAGACCAAGACGCTCTCAAGGAGAGACCAAGCCAAAGCAACAGACGATAGT 360  
Db 101 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 120

QY 361 CACCAGAGGTGACCGTTTCGATAGGGTGATCTCATTCAGATTCACCGGTGTGCT 420  
Db 121 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 140

QY 421 TTCTGGCTCTACACGACACGACACTGATGTGTGCTGCTCTTCTTCTTACTGACCCAAC 480  
Db 141 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 160

QY 481 AACCAACCAACACAGCTTGATCAGTTCCTCCAGGAGATTCAATTGGCTGGCAACACCGAG 540  
Db 161 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180

QY 541 CAAGAGTCTTAAAGTACACGACCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600  
Db 181 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 200

QY 601 AGCCCATACGCCCGCAAGTACGCTTAGACAGAGAGCGGTGATTTAGCCCTCGAGGA 660  
Db 201 SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 220

QY 661 CAGCACAGCCGACAGAACGAGGAGGACAAAGAAAGAAAGAGTGGAAACATCTTC 720  
Db 221 GlnHisSerArgArgGluArgAlaGlyGlnGluAsnGluGlyGlyAsnIlePhe 240

QY 721 AGCGGCTTCAGCCGGAGTTCCTGGAAACAGCCTTCAGGTTGACGACAGACAGATAGTG 780  
Db 241 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 260

QY 781 CAAAACCTTAGCGGAGCGGACCGAGCTCAAGAGAGGAGCGCATTTGACAGTGAGGGGA 840  
Db 261 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 280

QY 841 GGCTTCAGAAATCTTGAGCCAGATAGAAAGAGAGCGTCCGACGAAAGAGGAGATACGAT 900  
Db 281 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluGlyTyrAsp 300

QY 901 GAAGATGAATATGAATACGATGAAGAGGATAGAGGCGTGGCAGGGGAAGCAGAGGCGAG 960  
Db 301 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 320

QY 961 GGGATCGGATTGAAGAGACGATCTGACCGCAAGTCTGCTTAAAGAACATTTGGTAGAAC 1020  
Db 321 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 340

QY 1021 AGATCCCTGACATCTACACCTCAAGCTGGTTCACTCAAACTGCCCAACGATCTCAAC 1080  
Db 341 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 360

QY 1081 CTTCTAATACTTAGTGGCTTGACACCTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140  
Db 361 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 380

QY 1141 TTCTTTTGTGCTCAGTACCAACCAACGACACAGCATCATATATCGATTGAGGGGCGG 1200  
Db 381 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 400

QY 1201 GCTCAGCTGCACTCGTGACACGACCAACGACAGAGTGATCGACGAGGAGCTTCAAGAG 1260  
Db 401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 420

QY 1261 GGTACAGTCTGCTGTCGTCACAGAACTTCGCGCTCGCTGGAAGTCCCGAGGAGAAC 1320  
Db 421 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 440

QY 1321 TTCTGAATACGTGCGATTCAAGCAGACACTCAAGGCCAGCATAGCAACCTCGCGGTGAA 1380  
Db 441 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 460

QY 1381 AACTCCGTCATAGATAACCTCGCGAGGAGTGGTTGCAAAATTCATATGCGCTCCAAAGG 1440  
Db 461 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 480

QY 1441 GACGAGGCAAGGAGGAGCTTAAGAAACAACACCCCTTCAGTTCCTCGTCCACCGTCTCAG 1500  
Db 481 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln 500

QY 1501 CAGTCTCCGAGGCGCTGTGGCT 1521  
Db 501 GlnSerProArgAlaValAla 507

RESULT 3  
US-10-228-806-6  
; Sequence 6, Application US/10228806  
; Publication No. US20030049237A1  
; GENERAL INFORMATION:  
; APPLICANT: Bannon, et al.  
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions  
; TITLE OF INVENTION: to Allergy  
; FILE REFERENCE: 2002834-0043  
; CURRENT APPLICATION NUMBER: US/10/228,806  
; CURRENT FILING DATE: 2002-08-26  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 510  
; TYPE: PRT  
; ORGANISM: Arachis hypogaea  
US-10-228-806-6

Alignment Scores:  
Pred. No.: 1,55e-221 Length: 510  
Score: 2673.00 Matches: 507  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.33% Indels: 0  
DB: 14 Gaps: 0

US-10-728-051-3 (1-1524) x US-10-228-806-6 (1-510)

QY 1 CGGAGCAACCGGAGGAGAACGGCTGCCAGTTCAGCGCCTCAATGCGCAGACCTTGAC 60  
Db 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23

QY 61 AATCGCATTCGAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACCGAGGATTC 120  
Db 24 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 43

QY 121 GAATGCGCGCGCTGCGCCTCTCTCGCTTAGTCTCCGCGCAGAACCGCCTTCGTAGGCGCT 180  
Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 63

```
181 TTCTACTCCAATGCTCCCGAGGAGATCTTTCATCCAGCAAGAAAGGGATACATTTGGGTTG 240
Db
64 PheTyr-SerAsnAlaProGlnGluIlePhePheGlyArgGlyTyrPheGlyLeu 83
QY
241 ATATTCCCTGGTGTCTAGACACTATGAAGACCTCACAACAAGGTGCTGATCTCAG 300
Db
84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 103
QY
301 TCCCAAGACCAACAAGAGCTCTCCAGGAGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
Db
104 SerGlnArgProProArgArgLeuGlnGlyIleAspGlnSerGlnGlnArgAspSer 123
QY
361 CACCAGAAGGTGACCGGTTTCGATGAGGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
Db
124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY
421 TTCTGGCTCTACAGACCAAGACGACTGATGTTGGTGTCTTCTTACTGACACCAAC 480
Db
144 PheTrpLeuTyrAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn 163
QY
481 AACAAACACACAGCTTGATCAGTTCCCGAGGAGATTCAATTTGCTGGGAACCGGAG 540
Db
164 AsnAsnAspGlnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 183
QY
541 CAAGAGTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT 600
Db
184 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
QY
601 AGCCCATACAGCCCGCAAGTCCAGCTAGACAGAGAGCGTGAATTTAGCCCTCAGGA 660
Db
204 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 223
QY
661 CAGCAGACCGCAGAGAACAGCAGCAGCAGACAGAAAGAAACCAAGTGGAAACATCTTC 720
Db
224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGlnGlyAsnIlePhe 243
QY
721 AGCGGTTTACGCGGAGTCTTGGAAACAAGCCCTCCAGGTTCCAGCAGCAGACAGATAGTG 780
Db
244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspArgGlnIleVal 263
QY
781 CAAACCTTAGAGCGGAGCCGAGAGTGAAGAGAGGAGCCATTCTGACAGTGGGGA 840
Db
264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 283
QY
841 GGCTCAGAAATCTTGAGCCAGATAGAAAGACGTCGCCAGCAAGAAAGAAATACGAT 900
Db
284 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 303
QY
901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGAGCAGG 960
Db
304 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 323
QY
961 GGAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTTAAAGAACATTGTGTAGAAC 1020
Db
324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343
QY
1021 AGATCCCTGACATCTACAAACCTCTCAAGCTGGTTCACTCAAAACTGCCAACGATCTCAAC 1080
Db
344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
QY
1081 CTTCTTAATATCTTAGTGGCTTGGACCTAGTGTGTAATATGGAATCTCTACAGGAATGCA 1140
Db
364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgArgAla 383
QY
1141 TTGTTTGTCCCTCACTACACCAACCAAGCAGCAGCATCATATATCGATTGAGGGACGG 1200
Db
384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
QY
1201 GCTCAGCTGCAAGTCTGTGGACAGCAACCGCAACAGAGTGTACACGAGGAGCTTCAAGAG 1260
Db
404 AlaHisValGlnValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423
QY
1261 GGTCAAGTCTGTGTGGTGCACAGAACTTCGCGCTGCTGGAAGTCCAGAGCCAGAAC 1320
Db
424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
QY
1321 TTCGAATAGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAACTCCCGGTGAA 1380
Db
444 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
QY
1381 AACTCGTCTATAGATAAACCCTCGGAGGAGGTGGTTCAAATTCATATATGGCTCCAAAGG 1440
Db
464 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 483
QY
1441 GAGCAGGCAAGCAGCTTAAGAACAAACACCCCTTCAAGTCTTCCGTTCCACCGTCTCAG 1500
Db
484 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln 503
QY
1501 CAGTCTCCGAGGGCTGTGGCT 1521
Db
504 GlnSerProArgAlaValAla 510
RESULT 4
; Sequence 90, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 3
US-10-100-303A-90
Alignment Scores:
Pred. No.: 1-55e-221 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0
DB: 15 Gaps: 0
US-10-728-051-3 (1-1524) x US-10-100-303A-90 (1-510)
QY 1 CCGCAGCAACCGGAGAGAACCGGTGCCAGTTCCAGCGCTCAATGCGCAGAGACCTGAC 60
Db 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
QY 61 AATCCGATTGAATCAGAGCGGCTTACATTGAGACTTGGAAACCCCAACACAGAGATTC 120
Db 24 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 43
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAAGCCCTCTGTTAGGCT 180
Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 63
QY 181 TTCTACTCTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGATACATTTGGGTTG 240
Db 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGACCTCACAACAAGGTGCTGATCTCAG 300
Db 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 103
QY 301 TCCCAAGACCAACCAAGAGCTCTCCAGGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
Db 104 SerGlnArgProProArgArgLeuGlnGlyIleAspGlnSerGlnGlnArgAspSer 123
```

QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCACCGGTGTGCT 420  
Db 124 HieGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143  
QY 421 TTCTGGCTCTACAAACGACACGACACTGATGTTGTGCTGTTCTTCTTACTGACACCAAC 480  
Db 144 PheTrpLeuTyraAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163  
QY 481 AACAAACGACACGAGTTCATAGTTCCTCCAGGAGATTCATTTGGCTGGAAACACGGAG 540  
Db 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 183  
QY 541 CAAGAGTCTTAAGTACACGACCAACAAAGCAGCAAAAGCAGACGAGCAAGCTTACCATAT 600  
Db 184 GlnGluPheLeuArgTyrglnGlnGlnSerArgGlnSerArgArgSerLeuProTyf 203  
QY 601 AGCCATACACCGCGCAAGTTCAGCTCCTAGACAGAGAGCGTGAATTTAGCCCTCCAGGA 660  
Db 204 SerProTyfSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 223  
QY 661 CAGCACGCGCCAGAGAACGAGCAGGACAAGAGAGAAAGAAACGAAGGTGGAAACATCTTC 720  
Db 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 243  
QY 721 AGCGCTTCACGCGGAGTTCCTGGAAACAGACCTTCAGGTTGACGACACAGATAGTG 780  
Db 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263  
QY 781 CAAACCTTAAGAGCGCAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGGGGGA 840  
Db 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 283  
QY 841 GGCTTCAGAAATCTTGACCCAGATAGAAAGAGAGCGTCCGACGAAAGAGGAAATACGAT 900  
Db 284 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluTyfAsp 303  
QY 901 GAAGATGAATATGAATACGATGAAGAGATGAAGCGTGGCAGGAGGAGCAGGAGG 960  
Db 304 GluAspGluTyfGluTyfAspGluGluAspArgArgGlyArgGlySerArgGlyArg 323  
QY 961 GCGAATGGTATGAAGAGAGATCTGCACCGCAAGTCTAAAGACATTGCTAGAAAC 1020  
Db 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIysAsnIleGlyArgAsn 343  
QY 1021 AGATCCCTGACATCTACAAACCTTCAAGCTGGTTCACTAAAATCGCAACGATCTCAAC 1080  
Db 344 ArgSerProAspIleTyfAsnProGlnAlaGlySerLeuIysThrAlaAsnAspLeuAsn 363  
QY 1081 CTTCTAATACCTTAGGTGGCTTGACCTAGTGTGAATATGAAATCTCTACAGAAATGCA 1140  
Db 364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyfGlyAsnLeuTyfArgAsnAla 383  
QY 1141 TTGTTGTCTCCTCACTACAAACCCAGCAGCAGCATCATATATCGATTGAGGAGCGG 1200  
Db 384 LeuPheValAlaHisTyfAsnThrAsnAlaHisSerIleIleTyfArgLeuArgGlyArg 403  
QY 1201 GCTCAGCTGAAGTCGTGGACAGCAACGCAACAGAGTGATACACAGAGAGCTTCAAGAG 1260  
Db 404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyfAspGluGluLeuGlnGlu 423  
QY 1261 GGTCACTGTGTTGTGGCCACAGAACTTCGCCGTGCTGGAAAGTCCACAGAGCGAAGC 1320  
Db 424 GlyHisValLeuValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443  
QY 1321 TTGCAATACGTGGCTTCAAGACAGACTCAAGCCCGAGCTAGCCACCTCGCGGTGAA 1380  
Db 444 PheGluTyfValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463  
QY 1381 AACTCCGTCTAGATAACCTGCGCGGAGGTGGTTCGCAATTCATATGCTCCCAAGG 1440  
Db 464 AsnSerValIleAspAsnLeuProGluValValAlaAsnSerTyfGlyLeuGlnArg 483  
QY 1441 GAGCAGGCAAGCGCCTTAAGAAACAACACCCCTTCAAGTTCTTCTCCACCGTCTCAG 1500

Db 484 GluGlnAlaArgGlnLeuIysAsnAsnAsnProPheLysPheValProPserGln 503  
QY 1501 CAGTCTCCGAGGGCTGTGGCT 1521  
Db 504 GlnSerProArgAlaValAla 510  
RESULT 5  
US-09-731-221-79  
; Sequence 79, Application US/09731221  
; Patent No. US20020018778A1  
; GENERAL INFORMATION:  
; APPLICANT: Caplan, Michael  
; TITLE OF INVENTION: Passive Desensitization  
; FILE REFERENCE: 2002834-0103  
; CURRENT APPLICATION NUMBER: US/09/731,221  
; CURRENT FILING DATE: 2001-12-06  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 79  
; LENGTH: 526  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Arachis  
; OTHER INFORMATION: Hypogaea  
US-09-731-221-79  
Alignment Scores:  
Pred. No.: 1,64e-218 Length: 526  
Score: 2638.00 Matches: 502  
Percent Similarity: 99.01% Conservative: 0  
Best Local Similarity: 99.01% Mismatches: 5  
Query Match: 98.03% Indels: 0  
Gaps: 0  
DB: 9  
US-10-728-051-3 (1-1524) x US-09-731-221-79 (1-526)  
QY 1 CGGCACCAACCGGAGGAGAACGCGTCCAGTTCACGCGCTCAATGCGCAGACCTTGAC 60  
Db 5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24  
QY 61 AATCGCATTAATCAGAGGCGGTTCATTGAGACTTGGNAACCCCAACCAACAGGAGTTC 120  
Db 25 AsnArgIleGluSerGluGlyTyfIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44  
QY 121 GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCAACCGCTTCGTAGGCT 180  
Db 45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 64  
QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGAGGAGATCTTTGGGTTG 240  
Db 65 PheTyfSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyfPheGlyLeu 84  
QY 241 ATATTCCTGTTGTTCTAGACACTATGAAGAGCTCACACAAAGTCTGTCATCTCAG 300  
Db 85 IlePheProGlyCysProArgHisTyfGluGluProHisThrGlnGlyArgArgSerGln 104  
QY 301 TCCCAAGACACCAACAGAGCTCTCCAGGAGAGACCAACCAACAGCAGGATAGT 360  
Db 105 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124  
QY 361 CACCAAGAGTGCACCGTTTCGATGAGGTTGATCTCATTCAGTTCCTCCACCGTGTGCT 420  
Db 125 HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144  
QY 421 TTCTGCTCTTACAAACGACCAACGACACTGATGTTGTGCTGTTCTTCTTACTGACCAAC 480  
Db 145 PheTrpLeuTyfAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164  
QY 481 AACCAACGACACGAGCTTGAATTCATTCCTCCAGGAGATTCATTTGGCTGGAAACAGGAG 540  
Db 165 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 184

541 CAAGAGTTCTTAAGTACAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT 600  
185 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 204  
601 AGCCCATACAGCCGCAAGTACGCTAGCAGAGAGCGTCAATTTAGCCCTCCAGCA 660  
205 SerProTyrSerProGlnSerGlnProArgGlnGlnGluArgGluPheSerProArgGly 224  
661 CACACAGCCGACAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720  
225 GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluGluGluGlyGlyAsnIlePhe 244  
721 AGCGGTTTCAAGCGGAGTCTTGGAAACAAGCCCTCCAGGTTGACACAGACAGATAGTG 780  
245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspArgGlnIleVal 264  
781 CAAACCTTAGAGCCGAGCCGAGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGluGlyAlaIleValThrValArgGly 284  
841 GGCTCAGAACTTTGAGCCGACATAGAAAGAGAGAGCGTCCGACCAAGAAAGAAATACGAT 900  
285 GlyLeuArgAlaLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 304  
901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGAAGCAGAGCAGG 960  
305 GluAspGluTyrAlaTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 324  
961 GCGAATGGATTCAAGAGAGTCTCCAGCGCAAGTCTTAAAGAACATTTGGTAGAAGC 1020  
325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 344  
1021 AGATCCCTCAGATCTACAAACCTCAAGCTGGTTTCACTCAAACTGCCAAACGATCTCAAC 1080  
345 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364  
1081 CTTCTAATACTTAAAGTGGCTGAGCCTAGTGTGAATATGGAATCTTACAGGAATGCA 1140  
365 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 384  
1141 TTGTTTGTGCTCAGTACACCAACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1200  
385 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 404  
1201 GCTCAGCTGCAAGTGTGACAGCAACGCAAGCAGAGTGTACACAGAGAGCTTCAAGAG 1260  
405 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 424  
1261 GGTCAAGTGTGTTGTGTCACAGAACTTGCCTGCGTGGAAAGTCCAGAGCGAGAAC 1320  
425 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444  
1321 TTCGAATAGTGGCATTCAAGACAGACTCAAGGCCAGCAGATAGCCAACTCCGCGTGAA 1380  
445 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464  
1381 AACTCCGTCATAGATAACTGCCGAGAGGAGTGGTGCATAATTCATATGGCTTCCAAAGG 1440  
465 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 484  
1441 GAGCAGGCAAGGAGCTTAAGAACAAACACCCCTTCAAGTCTTCCGTTCCAGCTCCAG 1500  
485 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln 504  
1501 CAGTCTCCGAGGGCTGGCT 1521  
505 GlnSerProArgAlaValAla 511

RESULT 6  
US-10-899-551-6  
; Sequence 6, Application US/10899551  
; Publication No. US20050063994A1

GENERAL INFORMATION:  
; APPLICANT: Caplan, Michael J.  
; APPLICANT: Burks, A. Wesley  
; APPLICANT: Sampson, Hugh A.  
; APPLICANT: Howard, Sosin B.  
; APPLICANT: Bottomly, Kim H.  
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy  
; FILE REFERENCE: 2002834-0233  
; CURRENT APPLICATION NUMBER: US/10/899,551  
; CURRENT FILING DATE: 2004-07-26  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 6  
; LENGTH: 530  
; TYPE: PRT  
; ORGANISM: species Arachis hypogea  
US-10-899-551-6

Alignment Scores:  
Pred. No.: 4,83e-217 Length: 530  
Score: 2621.00 Matches: 500  
Percent Similarity: 98.62% Conservative: 0  
Best Local Similarity: 98.62% Mismatches: 7  
Query Match: 97.40% Indels: 0  
DB: 17 Gaps: 0

US-10-728-051-3 (1-1524) x US-10-899-551-6 (1-530)

QY 1 CGGCAGCAACCGAGGAGAGACGCGTCCAGCGCTCAATGCGCAGAGACTGAC 60  
Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43  
QY 61 AATCGCATTTGAATCAGAGCGGTTTACATGTGAGACTTGGAAACCCCAACACAGAGATTC 120  
Db 44 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnPhe 63  
QY 121 GAATGCGCGCGTCCGCTCTCTCGCTAGTCTCCGCGCAACGCCCTTCGTAGGCT 180  
Db 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 83  
QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGAGTACTTTGGGTTG 240  
Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103  
QY 241 ATATTCCCTGTTGTCTTAGACACTATGAAGAGCTTCACACAAAGTCTCGATCTCAG 300  
Db 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgSerGln 123  
QY 301 TCCCAAGACCAACCAAGAGCTCTCCAGGAGAGACCAAGCCCAACAGCAGAGATAGT 360  
Db 124 SerGlnArgProProIleArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 143  
QY 361 CACCAAGAGTGCACCGTTTCGATGAGGTGATCTCATTGACAGTTCACACCGCTGTTGCT 420  
Db 144 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 163  
QY 421 TTCTGGCTCTACACGACCCACGACTGATGTGTGTGTGTGTCTTCTTACTCACCACAC 480  
Db 164 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183  
QY 481 ACAACGACCAACAGCTTGATCAGTTCCTCCAGAGATTCATTTGGCTGGGACACGGAG 540  
Db 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 203  
QY 541 CAAGAGTTCCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600  
Db 204 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 223  
QY 601 AGCCCATACAGCCCGCAAGTACAGCTCAGCCTAGCAAGAGAGCGTGAATTTAGCCCTCGAGA 660  
Db 224 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 243  
QY 661 CAGCAGCCGCGAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720



```
Db 244 GlnHisSerArgArgGluArgAlaGlyGlnGluGlnGluGlnGlyGlyAsnIlePhe 263
Qy 721 AGCGGCTTACGCGCGAGTTCCTGGAAACAAGCCCTCCAGGTTGACGACAGACAGATAGTG 780
Db 264 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspArgGlnIleVal 283
Qy 781 CAAACCTAAGACGCGAGCGAGAGTGAAGAGAGGAGGACCATTTGACAGTGAAGGGA 840
Db 284 GlnAsnLeuArgGlyGluAsnGluSerGluGluGluGlyAlaIleValThrValArgGly 303
Qy 841 GGCCTCAGAACTCTTGAGCCAGATAGAAAGAGAGCTGCCGACGAAGAAGAGGAATAGAT 900
Db 304 GlyLeuArgIleLeuSerProAspArguysArgArgAlaAspGluGluGluGluArgP 323
Qy 901 GAAGATGAATATGATACGATGAAGAGGATAGAAGCGTGCGAGGGGAAGCAGAGGCGAG 960
Db 324 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 343
Qy 961 GGGATGCTATTGAAGAGACGATCTGCACCGCAAGTCTTAAAGAACATTTGTAGAAC 1020
Db 344 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIleGlyAsnIleGlyArg 363
Qy 1021 AGATCCCTCGACATCTACAACCCCTCAAGCTGTTTCACTCAAACTGCCAACGATCTCAAC 1080
Db 364 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuIleThrAlaAsnAspLeuAsn 383
Qy 1081 CTTCTAATACTAGTGGCTTGGACCTAGTCTGCTGAATATGAAATCTCTACAGGAATGCA 1140
Db 384 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 403
Qy 1141 TTGTTTCGCTCACTACACACCAACGACGACAGCATATATATATGATTCATGAGGGACGG 1200
Db 404 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 423
Qy 1201 GCTCACGTCAAGTCGTGGACGACGACGACGACGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 424 AlaHisValGlnValAlaAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 443
Qy 1261 GGTCACTGCTTGTGTGGCCACAGAACTTCGCGCTCGCTGGAAAGTCCCGAGCGAGAAC 1320
Db 444 GlyHisValLeuValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 463
Qy 1321 TTGGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCAACCTCGCGGTGAA 1380
Db 464 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 483
Qy 1381 AACTCCGCTCATAGATAACCTGCCGAGGAGTGTGCAAAATTCATATGCGCTCCCAAGG 1440
Db 484 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 503
Qy 1441 GAGCAGGCAAGCGCAGCTTAAGAAACAACAACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500
Db 504 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 523
Qy 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db 524 GlnSerProArgAlaValAla 530
```

## RESULT 7

```
US-10-899-551-58
; Sequence 58, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sossin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
```

```
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58
; LENGTH: 351
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-58
```

```
Alignment Scores:
Pred. No.: 1,3e-134 Length: 351
Score: 1665.00 Matches: 315
Percent Similarity: 97.83% Conservative: 0
Best Local Similarity: 97.83% Mismatches: 7
Query Match: 61.87% Indels: 0
DB: 17 Gaps: 0
```

US-10-728-051-3 (1-1524) x US-10-899-551-58 (1-351)

```
Qy 1 CGGCACCAACCGGAGGAGAACGCGTCCAGTTCACGGCTCAATCGCAGAGACCTGAC 60
Db 15 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 34
Qy 61 AATCGCATTTGAATCAGAGGCGGTTACATTGACACTTTGGAAACCCCAACCAACAGGAGTTC 120
Db 35 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 54
Qy 121 GAATGCGCGGCTCGCCCTCTCTCGTTAGTCTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 55 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 74
Qy 181 TTCTACTCCATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATCTTTGGGTTG 240
Db 75 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 94
Qy 241 ATATTCCTGGTGTCTTAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCAG 300
Db 95 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgGlnGln 114
Qy 301 TCCAAAGACCAACCAAGACGTCTTCAAGGAGAAAGACCAAGCCCAACCAACGAGATAGT 360
Db 115 SerGlnArgProProIleArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 134
Qy 361 CACCAAGAGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
Db 135 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 154
Qy 421 TTCTGCTCTACAAACCAACCAACGACACTGTATGTTGTTGCTGTTTCTTCTTACTGACCAAC 480
Db 155 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 174
Qy 481 AACCAACGACCAACGAGCTTGATCAGTTCCCGAGGAGATTCAATTTGGCTGGGAACACGAG 540
Db 175 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 194
Qy 541 CAAGAGTTCTTAAGGTATCCAGCAACAAAGCAGACAAAGCAGACGAGAAGCTTACCATAT 600
Db 195 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 214
Qy 601 AGCCCATACAGCCCGCAAGTACGCTTAGACAAGAAGAGCGTCAATTTAGCCCTCGAGGA 660
Db 215 SerProTyrSerProGlnSerGlnProArgGlnGluGluGluGluPheSerProArgGly 234
Qy 661 CAGCAGACCGCAGAGAACGAGCAGCAGACGAAGAAGAAGAAGAAAGCTGGAACATCTTC 720
Db 235 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGlyGlyAsnIlePhe 254
Qy 721 AGCGGCTTACGCGGAGTTCCTTGGAAACAAGCCCTTCAGGTTTCAGCAGACAGATAGTG 780
Db 255 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 274
Qy 781 CAAAACCTTAGAGCGGACCGCAGAGTGAAGAGAGGAGGCCATTTGACAGTGAAGGGA 840
Db 275 GlnAsnLeuArgGlyGluAsnGluSerGluGluGluGlyAlaIleValThrValArgGly 294
```



US-10-424-599-171702

; Sequence 171702, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 171702

; LENGTH: 481

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_126061C.1.1.pap

US-10-424-599-171702

Alignment Scores:

Pred. No.:	1.69e-120	Length:	481
Score:	1502.00	Matches:	297
Percent Similarity:	71.57%	Conservative:	68
Best Local Similarity:	58.24%	Mismatches:	91
Query Match:	55.82%	Indels:	54
DB:	15	Gaps:	7

US-10-728-051-3 (1-1524) x US-10-424-599-171702 (1-481)

Qy	1	CGGAGCAACCGGAGAGACGGGTGCGAGTTCAGCGCCTCAATGCGCAGACCTGAC	60
Db	23	ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp	42
Qy	61	AATCGCATTAAGTACAGAGGGGGTACATTAGACATTTGAAACCCCAACACCGAGTTC	120
Db	43	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe	62
Qy	121	GAATGCCCGCGTCGCTCTCTCGCTAGTCTCCGCGCAACCGCCTTCGTAGGCCT	180
Db	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro	82
Qy	181	TTCTACTCAATGTCCTCCAGGAGATCTTCATCAGCAAGAAAGGGGATCTTTGGTTG	240
Db	83	SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet	102
Qy	241	ATATTCCTGCTGCTAGACACTATGAAGAGCCTCACACAAAGTCTGTCGATCTCAG	300
Db	103	IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln	119
Qy	301	TCCCAAGACACCAAGACGTCCTCAAGGAGAGAACCAAGCCAAACAGCAACGAGATAGT	360
Db	120	SerSerArgPro-----GlnAspArg	126
Qy	361	CACCAAGAGTGACCGTTTGGATTCATTCAGTTCATTCAGTTCCTCCACCGGTGCT	420
Db	127	HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla	146
Qy	421	TTCTGGCTCTACACGACACGACACTGATGTTGTTGTTGTTCTTCTTACTGACACCAAC	480
Db	147	TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn	166
Qy	481	AACAACGACACACGATTTGATCAGTTCCTCCAGGAGATTCATTTGGTGGGAAACACGAG	540
Db	167	SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	186
Qy	541	CAAGAGTTCTTAAGGTACCAACAAAGCAGCAACAAAGCAGACGAGCAAGAGCTTACCATAT	600
Db	187	GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly-----	199
Qy	601	AGCCCATACGCCCGCAAGTACGCTTAGACAAGAGCGGTGAATTTAGCCCTCGAGGA	660

Db	200	-----		200
Qy	661	CAGCAGCGCGCAGAGAACGAGCAGGACAAAGAAAGAAAGAAAGAGTGTGAACATCTTC	720	
Db	201	ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlyGlySerIleLeu	220	
Qy	721	AGCGGCTTCCGCGGAGTTCCTGGACAAAGCCCTCCAGGTTGACGACACGACGATAGTG	780	
Db	221	SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal	239	
Qy	781	CAAAACCTAAGAGCGCAGACCGCAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGCAGGGA	840	
Db	240	ArgLysLeuGlnGlyGluAsnGluGluGluGlyAlaIleValThrValLysGly	259	
Qy	841	GGCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGAGCGTCCCGACGAAAGAGAG	891	
Db	260	GlyLeuSerValIleSerProThrGluGlnGlnArgProGluGluGluGlu	279	
Qy	892	GAATACGATCAAGATGAATATATACGATGAGAGGATAGAGGCGTGCAGGGGAGC	951	
Db	280	LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer	294	
Qy	952	AGAGGCGAGGGGGAATGTTAAGAGACGATCTGCACCCCAAGTCTAAAGAAAGAACATT	1011	
Db	295	Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle	311	
Qy	1012	GGTAGAAACAGATCCCTGACATCTACACCTCAAGCTCGTTCATCTCAAAATCGCCAAC	1071	
Db	312	GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr	331	
Qy	1072	GATCTCAACCTCTAATCTAGTGGCTTGGACCTAGTCTGAATATGAATCTCTAC	1131	
Db	332	SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg	351	
Qy	1132	AGAAATGCTTTGTCGCTCACTACACCAACCAAGCAGACACAGCATCATATATCGATTG	1191	
Db	352	LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu	371	
Qy	1192	AGGGGACGGGCTCAGCTGCAAGTCGTGACAGCAACGCGCAACAGAGTGTACGACGAGG	1251	
Db	372	AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu	391	
Qy	1252	CTTCAGAGGCTCAGCTGCTTGTGGTCCACAGAACTTCGCTCGTGGAAAGTCCCGAG	1311	
Db	392	LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln	411	
Qy	1312	AGCGAGAACTTCAATACGTGGCATTCACAGACAGATCAAGGCCACGACATAGCCAACTC	1371	
Db	412	SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu	431	
Qy	1372	GCCGTTGAAACTCCGTCATAGATAACCTCCGAGGAGGTGTTCGAAATTCATATGGC	1431	
Db	432	AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnThrPheAsn	451	
Qy	1432	CTCAAGAGGAGCAGCAGGAGCTTAAAGCAACACCCCTCAAGTCTTCGTTCCCA	1491	
Db	452	LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro	471	
Qy	1492	CCGCTCTCAGCAGTCCGAGGCGTGTGGCT	1521	
Db	472	ProLysGluSerGlnArgValValAla	481	

RESULT 10

US-10-425-114-43865

; Sequence 43865, Application US/10425114

; Publication No. US2004003488A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
NUMBER OF SEQ ID NOS: 73128

LENGTH: 488  
TYPE: PRT

ORGANISM: Glycine max

FEATURE:  
OTHER INFORMATION: Clone ID: 700645818\_FLI.pep

US-10-425-114-43865

Alignment Scores:  
Pred. No.: 1 7e-120 Length: 488  
Score: 1502.00 Matches: 297  
Percent Similarity: 71.57% Conservative: 68  
Best Local Similarity: 58.24% Mismatches: 91  
Query Match: 55.82% Indels: 54  
DB: 15 Gaps: 7

US-10-728-051-3 (1-1524) x US-10-425-114-43865 (1-488)

```
QY 1 CGGCAGCAACCGGAGGAGAACGGCTTCCAGTCCAGCGCTCAATCGCGAGAGACCTGAC 60
Db 30 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 49
QY 61 AATCGCATGTAATCAGAGGCGGTTCATTGAGACTTGGAAACCCCAACACGAGGATTC 120
Db 50 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 69
QY 121 GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCCTCGCGGCAACGCCCTTCGTAGCCT 180
Db 70 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 89
QY 181 TTCTACTCCAAATGCTCCAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGGTG 240
Db 90 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 109
QY 241 ATATTCCTGGTGTCTGACACTATGAGAGCCTCACACAAAGTGTGTCGATCTCAG 300
Db 110 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 126
QY 301 TCCCAAGACCAACAGAGCTCTCCAGGAGAGACCAAGCAACAGAGATAGT 360
Db 127 SerSerArgPro-----GlnAspArg 133
QY 361 CACCAAGAGTGCAACGGTTTCGATGAGGTGATCTCATTCAGTCCACCGGTGTGCT 420
Db 134 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 153
QY 421 TTCTGGCTTACACGACCAACGACACTGATGTTGTTGCTTCTTACTGACACCAAC 480
Db 154 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 173
QY 481 AACACGACCAACGAGTTCATGATTCCTCCAGGAGATTCATTTGCTGGGACAGGAG 540
Db 174 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 193
QY 541 CAAGAGTTCCTTAAGTACCAAGCAACAAAGCAGCAAGAGCAGACCAAGAGCTTACCATAT 600
Db 194 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 206
QY 601 AGCCCATACAGCCCGCAAGTACGCTCAGCAAGAGAGAGCGTGAATTTAGCCCTCAGGA 660
Db 207 -----Gly 207
QY 661 CAGCAGCGCAGAGACGAGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 208 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGluAsnGlnGlySerIleLeu 227
```

## RESULT 11

US-10-425-114-44047  
; Sequence 44047, Application US/10425114  
; Publication NO. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128

```
; SEQ ID NO 44047
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700763253_FLI.pep
US-10-425-114-44047

Alignment Scores:
Pred. No.: 1,7e-120 Length: 491
Score: 1502.00 Matches: 297
Percent Similarity: 71.57% Conservative: 68
Best Local Similarity: 58.24% Mismatches: 91
Query Match: 55.82% Indels: 54
DB: 15 Gaps: 7

US-10-728-051-3 (1-1524) x US-10-425-114-44047 (1-491)

QY 1 CGGCAGCAACGGAGGAGAGCGGTGCCAGTTCAGCGCCTCAATGCGCAGACCTGAC 60
Db 1 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 33 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 52
Db 33 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 AATCGCATTAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
Db 61 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 53 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnLysProPhe 72
Db 53 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 GAATGCGCGCGTCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGCCCT 180
Db 121 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 73 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuA-GatPro 92
Db 73 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 TTCTACTCCATGCTCCCGAGGAGATCTTCATCAGCAAGAGAGGGGATCTTGGTGTG 240
Db 181 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 93 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 112
Db 93 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 ATATTCCCTGCTGCTCTAGACACTATGAAGAGCTTCACACAAAGTGTGCTGATCTCAG 300
Db 241 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 113 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 129
Db 113 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 TCCCAAGACACCAACAGACGCTCTCCAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
Db 301 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 130 SerSerArgPro-----GlnAspArg 136
Db 130 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 CACGAGAGTGTGACCGTTCGATGAGGTGATCTGATTCAGTTCACCGGTGTGCT 420
Db 361 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 137 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 156
Db 137 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 421 TTCTGGCTCTACAGCACACGACACTGATGTTGCTGCTTCTTCTTACTGACACCAAC 480
Db 421 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 157 TyrTrpMetTyrAsnAsnGluAspThrProValAlaValSerLeuIleAspThrAsn 176
Db 157 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 481 AACACGACCAACACGCTTGATGATTCCTCCAGGAGATTCAATTTGGCTGGGAACACGGAG 540
Db 481 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 177 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 196
Db 177 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 541 CAGAGGTTTACAGTACAGCAACAAAGCAGCAAAAGCAGACGAAGAGCTTACCATAT 600
Db 541 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 197 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 209
Db 197 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 601 AGCCCATACGCCCGCAAGTACGCTTAGACAGAGAGCGGTGAATTTAGCCCTCGAGGA 660
Db 601 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 210 -----Gly 210
Db 210 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 661 CAGCAGCGCGCAGAACGAGCAGGAGCAAGAGAAAGAAACGAAGTGGAAACATCTTC 720
Db 661 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 211 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGluGlnGlyGlySerIleLeu 230
Db 211 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAAGCCCTTCAGGTTGACGACAGACAGATAGTG 780
Db 721 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 231 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 249
Db 231 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 781 CAAACCTTAAGCGCAGCAGCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGGAGGA 840
Db 781 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 250 ArgLysLeuGlnGlyGluAsnGluGluGluGlyGlyAlaIleValThrValLysGly 269
Db 250 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

## RESULT 12

```
US-10-425-114-43836
; Sequence 43836, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43836
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700673875_FLI.pep
US-10-425-114-43836
```

## Alignment Scores:

Pred. No.: 1.7e-120 Length: 492  
Score: 1502.00 Matches: 297  
Percent Similarity: 71.57% Conservative: 68  
Best Local Similarity: 58.24% Mismatches: 91  
Query Match: 55.82% Indels: 54  
DB: 15 Gaps: 7

US-10-728-051-3 (1-1524) x US-10-425-114-43836 (1-492)

```
QY 1 CGGCACACCGAGAGACGGCTCCAGGCTCAATCGGCAGAGACCTGAC 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 34 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 53
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 AATCGCATTAATCAGAGGCGGTTACATTGAGACTTGAACCCCAACACAGGAGTTC 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 54 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 73
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 GAATCGCGCGGCTCGCCCTCTCTCGCTTAGTCTCCGCCGCAACCCCTTCGTAGGCT 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 74 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 93
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 TTCTACTCCATCTCCCGAGGATCTTCATCCAGCAGGAGGAGGATCTTTGGGTTG 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 94 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 113
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 ATATTCCCTGTTGCTAGACACTATGAAGAGCCTCACACACAGGTGCTCGATCTCAG 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 130
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 TCCCAAGACCAACAGAGCTCTCCAGGAGAGACCAAGCCACAGCAACGAGATAGT 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 SerSerArgPro-----GlnAspArg 137
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 CACGAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCACCGGTGTGCT 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 138 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 157
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 421 TTCTGGCTTACACACGACGACGACTGATGTGTGTGCTTCTTCTTACTGACACCAAC 480
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 158 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 481 AACACGACACACGAGTTGATCTCCCGAGAGATTCATTTGCTGGGACACGGAG 540
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 178 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 197
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 541 CAAGAGTTCTTAGGTACACGACCAACAGCAGCAACAGCAGCAGCAGAGAGCTTACCAT 600
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 210
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 601 AGCCCATACAGCCGCAAGCTCAGCTAGACAGAGAGCGCTGATTTAGCCCTCGAGGA 660
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 211 -----Gly 211
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 661 CAGCAGACGCGCAGAGACGAGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 212 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlyGlySerIleLeu 231
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 721 AGCGGCTTACCGCGGAGTCTCTGGACACAGCCTTCAGGTTACACGACAGCAGATAGT 780
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 232 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 250
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 781 CAAACCTTAGAGCGCAGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 ArgLysLeuGlnGlyGluAsnGluGluGluGluGlyAlaIleValThrValLysGly 270
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 841 GGCTCAGATCTTGACCCA-----GATAGAAAGAGAGCTGCCGACGAGAGAGAG 891
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 271 GlyLeuSerValIleSerProThrGluGlnGlnGlnArgProGluGluGluGlu 290
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 892 GAATACGATCAAGATGAATATGATACGATGAGAGAGATAGAGGCGTGGCAGGGAAGC 951
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 291 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 305
QY 952 ACAGGCAGGGGAATGTTATGAAGACAGCATCTCCACCGCAAGTGCTAAAGAACATTT 1011
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 306 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 322
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1012 GGTAGAAACAGATCCCTCGACATCTACAACCTCAAGCTGGTTCACTCAAAATGCCAAC 1071
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 323 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 342
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1072 GATCTCAACCTTCTAATCTAGGTGGCTTGACCTAGTGTGAATATGGAATCTCTAC 1131
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 343 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheIleSerLeuArg 362
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1132 AGGAATGATCTTTGTTGCTCACTACAACCAACCAACGACACAGCATCATATCGATTG 1191
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 363 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 382
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1192 AGGACGCGGCTCAGTGCAGTCTGGACAGCAACGCAACAGAGTGTACACGAGGAG 1251
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 383 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 402
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1252 CTTCAAGAGGTCACGTGCTGTGTGTCACAGAACTTCGCCGCTCGTGAAAGTCCACAG 1311
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 403 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 422
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1312 AGCGAAGACTTCGAATACGTGCACTCAAGCAGACAGTCAAGGCCACGACATGCCAACCTC 1371
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 423 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 442
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1372 GCCGTGAAAACCTCCGTACATAGATAACCTGCCGAGGAGGTGGTGCAAATTCATATGGC 1431
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 443 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 462
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1432 CTCCAAGAGGACGACGAGGAGGAGCTTAAGAACACCAACCCCTTCAAGTCTTCTGTTCCA 1491
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 463 LeuArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 482
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1492 CGGCTCAGCAGCTCTCCGAGGCGCTGCGCT 1521
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 483 ProLysGluSerGlnArgArgValValAla 492
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 13
US-10-425-114-46294
; Sequence 46294, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46294
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700757865_FLI pep
US-10-425-114-46294
```

Alignment Scores:  
Pred. No.: 1.7e-120 Length: 492  
Score: 1502.00 Matches: 297  
Percent Similarity: 71.57% Conservative: 68  
Best Local Similarity: 58.24% Mismatches: 91  
Query Match: 55.82% Indels: 54



```
DB: 15 Gaps: 7
US-10-728-051-3 (1-1524) x US-10-425-114-46294 (1-492)
QY 1 CGGCAGCAACGGAGGAGACGGTTCAGCTTCAGCGCTCAATGCGCAGACCTGAC 60
DB 34 ArgGluGlnProGlnGlnAenGluCysGlnGlnArgLeuAenAlaLeuLysProAsp 53
QY 61 AATCGCATTAATCAGAGGGCGGTTCATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
DB 54 AsnArgileGluSerGluGlyGlyPheIleGluThrTrpAsnProAenAenLysProPhe 73
QY 121 GAATGCGCGGCTCGCCCTCTCTCGCTAGTCTCGCGCGCAACGCCCTTCGTAGGCT 180
DB 74 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAenArgAenAlaLeuArgGPro 93
QY 181 TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCAGCAAGAAAGGGATCTTTGGGTG 240
DB 94 SerTyrThrAenAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 113
QY 241 ATATTCCCTGGTTGCTTAGACACTATGAAGAGCTTCACACAAAGTCTGTCGATCTCAG 300
DB 114 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 130
QY 301 TCCCAAGACCAACAGACGCTCTCCAAGGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
DB 131 SerSerArgPro-----GlnAspArg 137
QY 361 CACCAAGAGTGCAACCGTTTCGATAGGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
DB 138 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 157
QY 421 TTTGCGCTTACCAACGACACGACACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
DB 158 TyrTrpMetTyrAenAenGluAepThrProValValAlaValSerLeuIleAepThrAen 177
QY 481 AACCAACGACACAGCTTGATGATTCCTCCAGGAGATTCATTTGGCTGGGAACAGGAG 540
DB 178 SerPheGlnAenGlnLeuAepGlnMetProArgArgPheTyrLeuAlaGlyAenGlnGlu 197
QY 541 CAAGAGTTCTTAAGGTACCAACAGCAACAAAGCAGCAAGCAGACGAGCAAGAGCTTACCATAT 600
DB 198 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly-----Gly 211
QY 601 AGCCCATACAGCCCGCAAGTACGCTTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
DB 211 -----Gly 211
QY 661 CAGCAGACCGCAGACAGCAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 212 ThrGlnSerGlnLysGlyTyrAsrGlnGlnGlnGluGluAenGluGlyGlySerIleLeu 231
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAGCCTTCAGGTTTCAGCAGACAGACAGATAGTG 780
DB 232 SerGlyPheAlaProGluPheLeuGluIleAlaPheValVal---AspArgGlnIleVal 250
QY 781 CAAAACTTAAGAGCGGAGACCGGAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 251 ArgLysLeuGlnGlyGluAenGluGluGluGluGlyAlaIleValThrValLysGly 270
QY 841 GGCTCAGATCTTGACCCCA-----GATAGAAGAGAGCTGCCGACGAGAGAGAGAGAG 891
DB 271 GlyLeuSerValIleSerProThrGluGlnGlnGlnArgProGluGluGluGlu 290
QY 892 GAATACCATCAAGATGAATATGATGATGAGAGGATAGAGGCGTGGCAGGGAAGC 951
DB 291 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 305
QY 952 AGAGGCGAGGGGAGTGTATTGAGAGACGATCTGCACCCCAAGTGTCTAAAGAGACATT 1011
DB 306 Arg-----AenGlyIleAepGluThrIleCysThrMetArgLeuArgHisAenIle 322
QY 1012 GGTAGAAACAGATCCCTGATCTACACCCCTCAAGCTGGTTCTCACTCAAAACTGCCAAC 1071
```

```
DB 323 GlyGlnThrSerSerProAspIlePheAenProGlnAlaGlySerIleThrThrAlaThr 342
QY 1072 GATCTCAACCTTCTAATACTTAGTGGCTTGGACCTTAGTGGCTGAATATGAAATCTCTAC 1131
DB 343 SerLeuAepPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 362
QY 1132 AGAATGCAATTTGTTCTGCTCCTACACAAACCAACGACGACATCATATATCGATTG 1191
DB 363 LysAenAlaMetPheValProHisTyrAenLeuAenAlaAenSerIleIleTyrAlaLeu 382
QY 1192 AGGGGACGGCTCACGTGCAAGTCGTGGACAGCAACGCAACAGAGTGTACGACGAGGAG 1251
DB 383 AsnGlyArgAlaLeuValGlnValValAenCysAenGlyGluArgValPheAspGlyGlu 402
QY 1252 CTTCAAGAGGGTCACTGCTTGTGGTCCACAGAACTTCGCGCTCGCTGGAAAGTCCCGAG 1311
DB 403 LeuGlnGluGlyGlnValLeuIleValProGlnAenPheAlaValAlaAArgSerGln 422
QY 1312 AGCGAAGACTTCGATACCTGCGCATTCACAGACAGACTCAAGGCCAGCATAGCCAACTC 1371
DB 423 SerAspAenPheGluTyrValSerPheLysThrAenAspArgProSerIleGlyAenLeu 442
QY 1372 GCGGTGAAACTCCGTCATAGTACCTGCGGAGGAGTGGTTCGAAATTCATATGGC 1431
DB 443 AlaGlyAlaAenSerLeuAenAlaLeuProGluGluValIleGlnGlnThrPheAen 462
QY 1432 CTCCAAGGAGGAGCAGGAGGAGCTTAAAGAACAAACACCCCTTCAAGTTCTTCGTTCCA 1491
DB 463 LeuArgGlnGlnAlaArgGlnValLysAenAenAenProPheSerPheLeuValPro 482
QY 1492 CGGTCTCAGCAGTCTCCGAGGCGTGGCT 1521
DB 483 ProLysGluSerGlnArgArgValValAla 492
```

## RESULT 14

```
US-10-425-114-71959
; Sequence 71959, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71959
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700676992_FLI pep
US-10-425-114-71959
```

```
Alignment Scores:
Pred. No.: 2,74e-120 Length: 466
Score: 1499.50 Matches: 297
Percent Similarity: 72.02% Conservative: 71
Best Local Similarity: 58.12% Mismatches: 94
Query Match: 55.72% Indels: 49
DB: 15 Gaps: 6
```

US-10-728-051-3 (1-1524) x US-10-425-114-71959 (1-466)

```
QY 1 CGGCAGCAACCGGAGGAGACGGTTCAGCTTCAGCGCTCAATGCGCAGACCTGAC 60
DB 1 ArgGluGlnAlaGlnGlnAenGluCysGlnIleGlnLysLeuAenAlaLeuLysProAsp 20
```



```
QY 61 AATCGCATTAATCAGAGCGCGTTACATTGAGACTTGAACCCCAACACAGGAGTTC 120
Db 21 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 40
QY 121 GAATGCGCGGCGTCCCTCTCTCGCTTAGTCTCCCGCGCAACGCCCTTCGTAGGCT 180
Db 41 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 60
QY 181 TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGACTTTGGTTG 240
Db 61 SerTyrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 80
QY 241 ATATTCCTCGTTGTCTTACACTATGAAGACCTTACACACAAGGTCTCGATCTCAG 300
Db 81 IlePheProGlyCysProSerThrTyrGlnGluProGlnGluSerGlnArgGlyArg 100
QY 301 TCCCAAGACCAACAGAGCTCCAGAGAGAACCAAGCCACAGCAACGAGATAGT 360
Db 101 SerGlnArgPro-----GlnAspArg 107
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
Db 108 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 127
QY 421 TTCTGGCTCTACCAACGACCACTGATGTTGTTGTTGTTGTTCTCTTACTGACACCAAC 480
Db 128 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 147
QY 481 AACACGACACACAGCTTGATGATTCCTCCAGAGAGATTCATTTGCTGGGCAACGAG 540
Db 148 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 167
QY 541 CAAGAGTCTTAAGTACAGCAACAAAGCAGACAAAGCAGACGAGCAAGAGCTTACCATAT 600
Db 168 GlnGluPheLeuLysTyrGlnGlnGln----- 176
QY 601 AGCCCATACAGCCGCAAGTACGCTTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGA 660
Db 177 -----GlnGlnGlySerGlnSerGlnLysGly 186
QY 661 CAGCAGCCGCGAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
Db 187 LysGln-----GlnGluGluAsnGluGlySerAsnIleLeu 199
QY 721 AGCGGTTACGCGGAGTTCCTGGAACAGCTTCAGGTTGACGACAGCAGCAGATAGTG 780
Db 200 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 218
QY 781 CAAACCTTAAGCGCGAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGGGGA 840
Db 219 ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 238
QY 841 GGCTCAGATCTTGACCCAGATAGAAAGAGAGCGTCCCGCAGAGAGAGGAAATACGAT 900
Db 239 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 258
QY 901 GAAGATGAA-----TATGAATACGATGAAGAGGATAGAACGCGTGGCAGGGA 948
Db 259 GluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 278
QY 949 AGCAGAGCGAGGGGAATGGTATTGAAGACGATCTGACCCGCAAGTGTCTAAAGAAC 1008
Db 279 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 295
QY 1009 ATTGGTAGAACAGATCCCTGATCTACAACTCAAGCTTCAAGTGTTCACCTCAAACTGCC 1068
Db 296 IleGlyGlnAsnSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 315
QY 1069 AACGATCTCAACCTTCTAATACTTAGTGGCTTGGACCTAGTGTGTAATATGGAATCTC 1128
Db 316 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 335
```

```
QY 1129 TACAGGAATGATTTGTTGCTCACTACAAACCAACGACACAGCATCATATATCGA 1188
Db 336 ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleTyrAla 355
QY 1189 TTGAGGGACGGGCTCAGCTGCAAGTCTGTCAGCAAGCAACGCAACAGAGTGTACACGAG 1248
Db 356 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 375
QY 1249 GAGCTTCAAGAGGCTCAGCTGCTTGTGTGTCACAGAACTTCGCGCTGCGTGGAAAGTCC 1308
Db 376 GluLeuGlnGluGlyGlyValLeuIleValProGlnAsnPheAlaValAlaAlaLysSer 395
QY 1309 CAGACGAGAACTTCGATAGTGGCTTCAAGCAGACTCAAGGCCAGCAGATAGCCCAAC 1368
Db 396 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 415
QY 1369 CTCGCGGTGAAACTCCGTCATAGATAAAGCTCCGCGAGGAGGTGGTTCGAAATTCATAT 1428
Db 416 LeuAlaGlyAlaAsnSerLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 435
QY 1429 GGCCTCCAAAGGAGCAGGCAAGCAGCTTAAAGAAACAACAACCCCTTCAAGTTCCTGTT 1488
Db 436 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 455
QY 1489 CCACGCTCTCAGCAGCTCCGAGGCTGTGGCT 1521
Db 456 ProProGlnGluSerGlnArgAlaValAla 466
```

RESULT 15  
US-10-425-114-71973  
; Sequence 71973, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 71973  
; LENGTH: 466  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700677092\_FLI.pep  
US-10-425-114-71973

Alignment Scores:  
Pred. No.: 2,74e-120 Length: 466  
Score: 1499.50 Matches: 297  
Percent Similarity: 72.02% Conservative: 71  
Best Local Similarity: 58.12% Mismatches: 94  
Query Match: 55.72% Indels: 49  
DB: 15 Gaps: 6

US-10-728-051-3 (1-1524) x US-10-425-114-71973 (1-466)

```
QY 1 CGGCAGCAACCGGAGAGAGAACCGGTGCCAGTTCAGCGCTCAATGCGCAGAGACCTGAC 60
Db 1 ArgGluGlnAlaGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 20
QY 61 AATCGCATTAATCAGAGCGGCTTACATTGAGACTTGAACCCCAACACAGGAGTTC 120
Db 21 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 40
QY 121 GAATGCGCGGCGTCCCTCTCTCGCTTAGTCTCCCGCGCAACGCCCTTCGTAGGCT 180
```

Db 41 GlnCysalaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 60  
Qy 181 TTCTACTCCAAAGTCTCCCGAGGAGATCTTTCATCCAGCAAGAGGGGATCTTTGGGTG 240  
Db 61 SerTyrThrAsnGlyProGlnGluIleTyrArgGlnGlyAsnGlyIlePheGlyMet 80  
Qy 241 ATATTCCCTGTTCTCTAGACACTATGAGAGCCTCACACAAAGTCTGCTCCATCTCAG 300  
Db 81 IlePheProGlyCysProSerThrTyrGlnGluProGlnGluSerGlnGlnArgGlyArg 100  
Qy 301 TCCCAAGAGACCAACAGACGCTCTCCAAGGAGAACCAAGCCAAAGCAACAGCATAGT 360  
Db 101 SerGlnArgPro-----GlnAspArg 107  
Qy 361 CACCAAGAGGTGACCGTTTCATGAGGGTGATCTCATTCAGTCTCCACCGGTGTGCT 420  
Db 108 HisGlnIysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 127  
Qy 421 TTCTGGCTCTACACGACACGACGACACTGATGTTCTGCTGTTCTTCTTACTGACACCAAC 480  
Db 128 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 147  
Qy 481 AACCAACGACCAACAGCTTGTAGTCTCCCGAGGAGATTCAATTTGGCTGGGAAACACCGAG 540  
Db 148 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 167  
Qy 541 CAAGAGTCTTAAGGTACCGAGCAACAAAGCAGCAACAAAGCAGCAACAAAGCTTACCATAT 600  
Db 168 GlnGluPheLeuIysTyrGlnGln-----176  
Qy 601 AGCCCATACGCCCGCAAGTCCAGCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGA 660  
Db 177 -----GlnGlnGlySerGlnSerGlnLysGly 186  
Qy 661 CAGCACAGCCCGCAGAGAACGAGCAGCAAGAGAAAGAAAGAGGTGGAACATCTTC 720  
Db 187 LysGln-----GlnGluGluAsnGluGlySerAsnIleLeu 199  
Qy 721 AGCGGCTTACGCGGAGTCTCTGGAACAAGCCTTCCAGGTTGACGACAGACAGATAGTG 780  
Db 200 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 218  
Qy 781 CAACAACCTAGAGCGGAGCCGAGAGTGAGAGAGGAGGAGCCATTTGACAGTGCAGGGGA 840  
Db 219 ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 238  
Qy 841 GGCCTCAGAATCTTGAGCCCAAGATAGAAAGAGCTGCCGACGAAGAGAGGAATACGAT 900  
Db 239 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 258  
Qy 901 GAAGATGAA-----TATGAATACATGAAGAGGATAGAACGCGTGGCAGGGGA 948  
Db 259 GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 278  
Qy 949 AGCAGAGCGGCGGAGTATTGTAAGAGACGATCTGCACCCGCAAGTGTCTAAAGAAC 1008  
Db 279 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 295  
Qy 1009 ATTGGTAGAAACAGATCCCTGCATCTACAACCTCTCAAGCTGTTCACTCAAAAGTCC 1068  
Db 296 IleGlyGlnAsnSerSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 315  
Qy 1069 AACGATCTCAACCTTCTAATCTTAGGTGGCTCGACCTAGTGTGAATATGGAATCTC 1128  
Db 316 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 335  
Qy 1129 TACAGGATGCAATGTTTGTGCTCACTACACACCAACGACACACATCATATATCGA 1188  
Db 336 ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 355  
Qy 1189 TTGAGGCGCGGCTCAGCTCAGTCTGAGCAACGCGCAACGCGCATGAGTGTACGAGAG 1248  
Db 356 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 375

Qy 1249 GAGCTTCAAGAGGGTCACGTGCTTGTGTGTCGCCACAGAACTTCGCGTCTGGAAGTCC 1308  
Db 376 GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaLysSer 395  
Qy 1309 CAGAGCGAGAACTTTCGAATACGTGGCATTCACAGACAGACTCAAGGCCACGATAGCCAAC 1368  
Db 396 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 415  
Qy 1369 CTCGCCGCTGAAAACTCCGTCATAGATAACCTCCGAGGAGGTGGTTGCAAAATTCATAT 1428  
Db 416 LeuAlaGlyAlaAsnSerLeuLeuAlaLeuProGluGluValIleGlnHisThrPhe 435  
Qy 1429 GGCCTCAAGAGGAGCAGGCAAGCAGCTTAAAGCAACCAACCCCTTCAAGTCTTCGTT 1488  
Db 436 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnPropheSerPheLeuVal 455  
Qy 1489 CCACCGCTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521  
Db 456 ProProGlnGluSerGlnArgAlaValAla 466

Search completed: April 8, 2005, 06:49:19  
Job time : 200.941 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 8, 2005, 03:43:40 ; Search time 203.83 Seconds

(without alignments)  
5783.472 Million cell updates/sec

Title: US-10-728-051-3

Perfect score: 2691

Sequence: 1 cggcagcaaccggaggagaa.....ctccgagggtgtgcttaa 1524

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10728051/runat\_07042005\_125044\_14041/app.query.fasta\_1.4757  
-DB=A\_Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0 -1 -LOOPTCT=0 -LOOPTXT=0  
-UNITS=bits -START=1 -END=1 -WATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pcp -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10728051@cgn 1 1 672 @runat\_07042005\_125044\_14041 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2673	99.3	510	6	ABU52484
2	2673	99.3	510	7	ADG27544
3	2667	99.1	510	4	AAU04708
4	2665	99.0	507	8	ADM12139
5	2663	99.0	510	2	AAI15246
6	2652	98.6	512	2	AAI40912
7	2638	98.0	526	3	AAB33601
8	2638	98.0	526	4	AAU04711
9	2638	98.0	526	4	AAU05036
10	2615	97.2	526	8	ADO38357

11	1502	55.8	481	5	ABG71266	Abg71266 Glycine m
12	1502	55.8	481	7	ADH89253	Adh89253 G. max gl
13	1502	55.8	481	7	ADL90187	Adl90187 Soybean g
14	1502	55.8	481	8	ADG43988	Adg43988 G. max gl
15	1499.5	55.7	485	5	ABG71265	Abg71265 Glycine m
16	1499.5	55.7	485	7	ADH89247	Adh89247 G. max gl
17	1499.5	55.7	485	7	ADL90186	Adl90186 Soybean g
18	1499.5	55.7	485	8	ADG43982	Adg43982 G. max gl
19	1496	55.6	495	3	AAI80994	Aai80994 Soybean g
20	1496	55.6	495	5	ABG71264	Abg71264 Glycine m
21	1496	55.6	495	7	ADH89245	Adh89245 G. max gl
22	1496	55.6	495	7	ADL90168	Adl90168 Soybean g
23	1496	55.6	495	8	ADG43980	Adg43980 G. max gl
24	1496	55.6	511	7	ADL90190	Adl90190 Soybean g
25	1492.5	55.5	485	6	ABU52502	Abu52502 Soybean g
26	1492.5	55.5	485	7	ADG27563	Adg27563 Soybean g
27	1488	55.3	484	2	AAI40949	Aai40949 Soybean g
28	1466	54.5	495	4	AAE10365	Aae10365 Soybean g
29	1034.5	38.4	457	8	ADO43062	Ado43062 Caehew nu
30	995.5	37.0	484	1	AAI71081	Aai71081 Sequence
31	986	36.6	517	5	ABG71267	Abg71267 Glycine m
32	986	36.6	517	7	ADL90189	Adl90189 Soybean g
33	973.5	36.2	562	7	ADH89249	Adh89249 G. max gl
34	973.5	36.2	562	7	ADL90188	Adl90188 Soybean g
35	973.5	36.2	562	8	ADG43984	Adg43984 G. max gl
36	946.5	35.2	562	5	ABG71268	Abg71268 Glycine m
37	941	35.0	561	1	AAI61363	Aai61363 Soybean g
38	926.5	34.4	516	7	ADH89251	Adh89251 G. max gl
39	926.5	34.4	516	8	ADG43986	Adg43986 G. max gl
40	914	34.0	185	2	AAI40951	Aai40951 Ara h 3 a
41	901.5	33.5	516	1	AAI61362	Aai61362 Soybean g
42	891	33.1	291	8	ADO60333	Ado60333 Cholester
43	845.5	31.4	499	7	ADC08275	Adc08275 Rice prot
44	841	31.3	499	7	ADC08135	Adc08135 Rice prot
45	841	31.3	499	7	ADC08279	Adc08279 Rice prot

ALIGNMENTS

RESULT 1  
ABU52484  
ID ABU52484 standard; protein; 510 AA.  
AC ABU52484;  
XX XX  
DT 10-MAR-2003 (first entry)  
XX XX  
DE Peanut Ara h3 protein sequence.  
XX XX  
KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;  
KW anaphylactic food allergen; anti-allergenic; vaccine; wound healing.  
XX XX  
OS Arachis hypogaea.  
XX XX  
DN WO200274250-A2.  
XX XX  
PD 26-SEP-2002.  
XX XX  
PF 18-MAR-2002; 2002WO-US009108.  
XX XX  
PR 16-MAR-2001; 2001US-0276822P.  
PR 18-MAR-2002; 2002US-00276822.  
XX XX

(PANA-) PANACEA PHARM.  
XX XX  
PA ABU52484 Peanut Ar  
XX XX  
PI Adg27544 Peanut al  
XX XX  
PI AAU04708 Anaphylac  
PI Adm12139 Arachis h  
XX XX  
XX Aai15246 Peanut al  
XX XX  
XX Aai40912 Ara h 3 a  
XX XX  
XX Aab33601 Modified  
XX XX  
XX Aau04711 Modified  
XX XX  
XX Aau05036 Modified  
XX XX  
XX Adu38357 Peanut al

Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;  
Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;  
Rabjohn PA, Shain DS, Stanley JS;  
WPI; 2003-018765/01.  
N-PSDB; ABX70612.  
New modified anaphylactic food allergen, useful for preventing or

treating allergic reactions associated with e.g. anaphylactic allergens.

Claim 27; Fig 68B; 300pp; English.

The invention relates to a modified anaphylactic food allergen has an amino acid sequence that is substantially identical to that of natural anaphylactic food allergen, except for a cysteine residue that has been modified so that it cannot participate in the disulfide bond. The modification may also comprise mutation of the IgE binding sites to reduce allergenicity. Also included are: (1) a method of making a modified anaphylactic food allergen; (2) a nucleotide molecule encoding for causing a site specific mutation in the modified anaphylactic food allergen; (3) a transgenic plant or animal expressing the modified anaphylactic food allergen; (4) a method of treating an individual by reducing the clinical response to a natural anaphylactic food allergen; and an isolated fragment of peanut allergen Ara h 1. The modified anaphylactic food allergen is useful for preventing or treating allergic reactions associated with any natural allergen such as food, insect, rubber or preferably anaphylactic allergens. It is also useful for treating wounds in mammals such as bovine, canine, feline, caprine, ovine, porcine, murine or equine species. The present sequence is a peanut allergen (e.g. Ara h1, h2 or h3)

XX Sequence 510 AA;

#### Alignment Scores:

Pred. No.: 3,77e-236 Length: 510  
Score: 2673.00 Matches: 507  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.33% Indels: 0  
DB: 6 Gaps: 0

US-10-728-051-3 (1-1524) x ABUS2484 (1-510)

QY	1	CGCAGCAACCGAGGAGACGGCTGCCAGTTCACGGGCTCAATCGCGAGACCTGAC	60
DB	4	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	23
QY	61	AATCGCATTAATCAGAGGCGGTTCATTGACACTTGGAAACCCCAACACAGGAGTTC	120
DB	24	AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnGlnGluPhe	43
QY	121	GAATGCGCGCGCTCGCTCTCTCGCTTAGTCTCTCGCGCAACGCCCTTCGTAGCCT	180
DB	44	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro	63
QY	181	TTCTACTCCATGCTCCCGAGGAGATCTTCATCCAGCAGGAGGGGATCTTTGGTTC	240
DB	64	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	83
QY	241	ATATTCCCTGTTGCTCTAGACACTATGAAGAGCCTCACACAAAGTCTGCTCATCTCAG	300
DB	84	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln	103
QY	301	TCCCAAGACCAACAGACTCTCCAGGAGAGACAAAGCCAAACAGCAACGAGATAGT	360
DB	104	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	123
QY	361	CACGAGAGTGACCGTTCGATGAGGTGATCTCATTCAGTCTCCACCGGTGTGCT	420
DB	124	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	143
QY	421	TTCTGCTCTACACCAACCAACGACACTGATGTTGTTGCTGTTTCTTCTTACTGACCAAC	480
DB	144	PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	163
QY	481	AACAACCAACCAACGCTTGATTCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACAGGAG	540
DB	164	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	183
QY	541	CAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT	600

#### RESULT 2

ADG27544

ID ADG27544 standard; protein; 510 AA.

XX

AC ADG27544;

XX

DT 26-FEB-2004 (first entry)

DB	184	GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr	203
QY	601	AGCCCATACAGCCCAAAAGTCAGCTAGACAAGAAGACGCTGAATTTAGCCCTCGAGA	660
DB	204	SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly	223
QY	661	CAGCAGCCGAGAGACGAGCAGGACAGAGAAAGAAAGAAAGAGGTTGGAACATCTTC	720
DB	224	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe	243
QY	721	AGCGCTTCACGCCGAGTCTCTGCAACAAGCTTCCAGTTTGACGACAGACAGATAGTG	780
DB	244	SerGlyPheThrProGluPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal	263
QY	781	CAAAACCTTAAGAGGCGACAGCAGAGTGAAGAAGAGGAGCCATTGTGACAGTGGGGA	840
DB	264	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	283
QY	841	GGCTCAGAACTTTCAGCCAGATGAAGAAGACAGCTGCCGACGAGAGAGAAATACGAT	900
DB	284	GlyLeuArgIleLeuSerProAspArgIysArgArgAlaAspGluGluGluTyrAsp	303
QY	901	GAAGATGAATATGATACGATGAAGAGGATGAAGCGCTGGCAGGGAAGCAGAGCGCAG	960
DB	304	GluAspGluTyrGluTyrAspGluAspArgArgGlyArgGlySerArgGlyArg	323
QY	961	GGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAACAATTCGTAGAAC	1020
DB	324	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	343
QY	1021	AGATCCCTGACATCTACAACTCAGCTGCTTCACTCAAACTGCCAACCATCTCAAC	1080
DB	344	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	363
QY	1081	CTTCTAATTAATCTAGTGGCTTGCACCTAGTCTGAATATGGAATCTCTACAGGAATGCA	1140
DB	364	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	383
QY	1141	TTGTTTGTGCTCCTACCAACACCAACGACACAGCATCATATATCGATTGAGGGACGG	1200
DB	384	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	403
QY	1201	GCTCAGTCCAGTCTGTCGACAGCAACGACAGAGTGTACGACGAGGAGCTTCAAGAG	1260
DB	404	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluGlnGlu	423
QY	1261	GGTCACTGCTTGTGTGTCACAGAACTTCGCGCTGCTGGAAGTCCGACAGCGAGAAC	1320
DB	424	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	443
QY	1321	TTCGAATACGTGGCATTTCAAGACAGACTCAAGGCCGAGCATAGCCAACTCCCGGTGAA	1380
DB	444	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	463
QY	1381	AACTCCGTCTATAGATAAAGTCCGGGAGGAGTGTTCGAATTCATATGCTCCCAAGG	1440
DB	464	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	483
QY	1441	GAGCAGCGAAGCGAGCTTAAGAACAACACCCCTTCAAGTCTTCTGTCACCGCTCTAG	1500
DB	484	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProSerGln	503
QY	1501	CAGTCTCCAGGGCTGTGGCT	1521
DB	504	GlnSerProArgAlaValAla	510

XX Peanut allergen Ara h3.  
 XX  
 XX  
 KW Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycinin A2B1a; Jug n1;  
 KW antiallergic; vulnerary; anaphylactic food allergen; IgE; allergy; wound.  
 XX  
 XX Arachis hypogaea.  
 XX  
 XX  
 XX US2003202980-A1.  
 XX  
 XX 30-OCT-2003.  
 XX  
 XX 18-MAR-2002; 2002US-00100303.  
 XX  
 XX 29-DEC-1995; 95US-0009455P.  
 PR 23-SEP-1996; 96US-00717933.  
 PR 31-JAN-1998; 98US-0073283P.  
 PR 13-FEB-1998; 98US-0074590P.  
 PR 13-FEB-1998; 98US-0074624P.  
 PR 13-FEB-1998; 98US-0074633P.  
 PR 29-JUN-1998; 98US-00106872.  
 PR 27-AUG-1998; 98US-00141220.  
 PR 13-NOV-1998; 98US-00191593.  
 PR 29-JAN-1999; 99US-00240557.  
 PR 29-JAN-1999; 99US-00241101.  
 PR 11-FEB-1999; 99US-00248673.  
 PR 11-FEB-1999; 99US-00248674.  
 PR 02-MAR-1999; 99US-0122450P.  
 PR 02-MAR-1999; 99US-0122452P.  
 PR 02-MAR-1999; 99US-0122560P.  
 PR 02-MAR-1999; 99US-0122565P.  
 PR 02-MAR-1999; 99US-0122566P.  
 PR 11-MAR-1999; 99US-00267719.  
 PR 28-JAN-2000; 2000US-00494096.  
 PR 16-MAR-2001; 2001US-0276822P.  
 XX  
 XX (CAPL/) CAPLAN M J.  
 PA (SOSI/) SOSIN H B.  
 PA (SAMP/) SAMPSON H.  
 PA (BANN/) BANNON G A.  
 PA (BURK/) BURKS A W.  
 PA (COCK/) COCKRELL G.  
 PA (COMP/) COMPADRE C M.  
 PA (CONN/) CONNAUGHTON C.  
 PA (HELM/) HELM R M.  
 PA (KING/) KING N E.  
 PA (KOPP/) KOPPER R A.  
 PA (MALE/) MALEKI S J.  
 PA (RABJ/) RABJOHN P A.  
 PA (SHIN/) SHIN D S.  
 PA (STAN/) STANLEY J S.  
 XX  
 XX Caplan MJ, Sosin HB, Sampson H, Bannon GA, Burks AW, Cockrell G,  
 PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;  
 PI Rabjohn PA, Shin DS, Stanley JS;  
 XX  
 XX WPI; 2003-875632/81.  
 DR N-PSDB; ADG27543.  
 DR  
 XX  
 XX Example 16; SEQ ID NO 89; 194pp; English.  
 XX  
 XX The invention relates to a modified anaphylactic food allergen whose  
 CC amino acid sequence is substantially identical to that of a natural  
 CC anaphylactic food allergen. The natural anaphylactic food allergen  
 CC includes at least one cysteine residue that participates in a disulphide  
 CC bond when the natural anaphylactic food allergen is in its native  
 CC conformation, except that the cysteine residue has been modified so that  
 CC it cannot participate in the disulphide bond. Also included are a method  
 CC of making a modified anaphylactic food allergen, a nucleotide molecule

CC encoding a modified anaphylactic food allergen defined above, a  
 CC nucleotide molecule for causing a site specific mutation in a gene  
 CC encoding a natural anaphylactic food allergen, a transgenic plant or  
 CC animal expressing a modified anaphylactic food allergen defined above, a  
 CC method of treating an individual by reducing the clinical response to a  
 CC natural anaphylactic food allergen by administering a modified  
 CC anaphylactic food allergen and an isolated fragment of peanut allergen  
 CC Ara h 1, comprising at least 10 consecutive amino acids of ADG27464 or  
 CC ADG27465. About 10-17% of the amino acids have been modified in at least  
 CC one IgE epitope or all the IgE epitopes recognised when the natural  
 CC anaphylactic food allergen is contacted with serum IgE from individual(s)  
 CC allergic to the natural anaphylactic food allergen. The invention  
 CC discloses Peanut allergens Ara h1, Ara h2, Ara h3 (and their encoding  
 CC cDNAs), Soybean Glycinin A2B1a and IgE-binding epitopes of the English  
 CC walnut allergen Jug n1. The modified anaphylactic food allergen can be  
 CC used for treating allergic reactions or wounds. The present sequence  
 CC represents a Peanut allergen of the invention (or its fragment).  
 XX  
 XX SQ Sequence 510 AA;  
 Alignment Scores:  
 Pred. No.: 3,77e-236 Length: 510  
 Score: 2673.00 Matches: 507  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.33% Indels: 0  
 DB: 7 Gaps: 0  
 US-10-728-051-3 (1-1524) x ADG27544 (1-510)  
 QY 1 CGGCAGCAACCGGAGAGAACGGTCCAGTTCACGGCGCTCAATGGCGAGACCTGAC 60  
 DB 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23  
 QY 61 AATCGCATTTGAATCAGAGGCGGTTTACATTGACACTTGGAAACCCCAACACAGGAGCTTC 120  
 DB 24 AsnArgIleGluSerGluGlyTyrlleGluThrTrpAsnProAsnGlnGluPhe 43  
 QY 121 GAATGCGCGGGCGTCCCGCTCTCTCGCTTAGTCTCTCCGCGCAACGCCCTTCGTAGSCCT 180  
 DB 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 63  
 QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATCTTTGGTTG 240  
 DB 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83  
 QY 241 ATATTCCTCGTTGTCTTAGACACTATGAGAGCCCTCACACAAAGTCGTCTGATCTCAG 300  
 DB 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 103  
 QY 301 TCCCAAGAGACCACCAAGAGCGTCTCCAAAGGAGAGAGACCAAGCCACAGCAACGAGATAGT 360  
 DB 104 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123  
 QY 361 CACCAGAGGTGACCGTTTCGATGAGGTGATCTCATTCGACGTTCACCGGTGTTGCT 420  
 DB 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143  
 QY 421 TTCTGCTCTACACAGCACCAGCACACTGATGTTGTGTGTTCTCTTACTGACACCAAC 480  
 DB 144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163  
 QY 481 AACACAGCAACACCGAGCTTGATCTCCCGAGGAGATTCAAATTTGGTGGGAACACCGAG 540  
 DB 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183  
 QY 541 CAAAGATTTTAAGGTACAGCAACAAAGCAGACAAAGCAGACGAGAGAGCTTACCATAT 600  
 DB 184 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203  
 QY 601 AGCCCATACAGCCGCAAGTACGCTAGACAGAGAGCGTCAATTTAGCCCTCCGAGGA 660  
 DB 204 SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 223





Db 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlyArgGlyTyrPheGlyLeu 83  
Qy 241 ATATTCCCTGGTGTCTCTAGACACTATGAAGAGCCTCACACAGAGTGTGTCGATCTCAG 300  
Db 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 103  
Qy 301 TCCCAAGACCAACAGAGCTCTCCAGAGGAGAGCAACCAAGCCAAAGCAACGAGATAGT 360  
Db 104 SerGlnArgProProArgArgLeuGlnGlyAspGlnSerGlnGlnArgAspSer 123  
Qy 361 CACCAGAGGTGACCGCTTTCAGTGGAGGTGATCTCATTCGAGTTCACCGGTGTTGCT 420  
Db 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143  
Qy 421 TTCTGGCTCTACACGACCAACGACACTGATGTTGTCTGCTGTTCTTCTTACTGACCAAC 480  
Db 144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163  
Qy 481 AACACCAACCAACGAGCTTGATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACGAG 540  
Db 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183  
Qy 541 CAAGAGTCTTAAAGGTACCAAGCAACAAAGCAGACAAAGCAGACCAAGCAAGCTTACCATAT 600  
Db 184 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203  
Qy 601 AGCCATATACCCCGCAAGTACGCTAGACCAAGAGCGTGAATTTAGCCCTCGAGGA 660  
Db 204 SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 223  
Qy 661 CAGCAGCGCGAGAGAACGAGCAGAGCAGAGCAAGAGAAACGAGGTGGAACATCTTC 720  
Db 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGlnGluAsnGluGlyGlyAsnIlePhe 243  
Qy 721 AGCGGCTTACCGCGGAGTCTCGAACACGCTTCCAGTGTGACGACAGCAGATAGTG 780  
Db 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspArgGlnIleVal 263  
Qy 781 CAAAACCTAAGAGCGAGACCGAGAGTGAAGAGAGGAGCGCATTTGTACAGTGAAGGGA 840  
Db 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGlyAlaIleValThrValArgGly 283  
Qy 841 GGCTCAGATCTTGACCCAGATAGAAAGAGAGCTGCCAGCAAGAGAGGATACGAT 900  
Db 284 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluGluTyrAsp 303  
Qy 901 GAGATCAATATGATACGATGAGAGAGATAGAGCGGTGCGAGGGAACGAGGCGAG 960  
Db 304 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyThr 323  
Qy 961 GGGAATGGTATTGAAGAGACGATCTGCCCGCAAGTCTAAAGAAACATTTGGTAGAAAC 1020  
Db 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343  
Qy 1021 AGATCCCTCAGATCTCAACCCCTCAAGCTGTTCTCACTCAAACTGCCAACGATCTCAAC 1080  
Db 344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363  
Qy 1081 CTTCTATACTAGTGGCTGGACCTAGTGTGCTGAATATGGAATCTCTACAGGAATGCA 1140  
Db 364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383  
Qy 1141 TTGTTTCTCGCTCACTACACCAACCAACGACATCATATATCATTCATTGAGGGGACGG 1200  
Db 384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403  
Qy 1201 GCTCAGTGCAGTGTGGAGCAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260  
Db 404 AlaHisValGlnValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423  
Qy 1261 GGTACAGTGTGTGGTGCACAGAACTTGGCCGTCTCGTGAAGTCCGAGGCGAGAAC 1320

Db 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443  
Qy 1321 TTCCAATACGTGGCATTTCAAGACAGACTCAAGGCCAGCATAGCAACCTCGCCGGTGA 1380  
Db 444 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463  
Qy 1381 AACTCCGTCTAGATACCTGCGGAGGAGGTGTTGCAAAATTCATATGSCCTCCAAAGG 1440  
Db 464 AsnSerValIleAspAsnLeuProGluGluValAlaAsnSerTyrGlyLeuGlnArg 483  
Qy 1441 GAGCAGCAAGGAGCGCTTAAGAACAAACCCCTTCAAGTCTTCTGTTTCCACCGCTCAG 1500  
Db 484 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 503  
Qy 1501 CAGTCTCCGAGGCGTGGCT 1521  
Db 504 GlnSerProArgAlaValAla 510  
RESULT 4  
ADM12139  
ID ADM12139 standard; protein; 507 AA.  
XX  
AC ADM12139;  
XX  
DX 20-MAY-2004 (first entry)  
XX  
DE Arachis hypogaea 3 (Ara h3) protein.  
XX  
KW antigen presentation enhancing hybrid polypeptide; mammalian II-Key;  
KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;  
KW antiarthritic; neuroprotective; dermatological; immunosuppressive;  
KW antiinflammatory; antidiabetic; antithyroid; immune;  
KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;  
KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;  
KW scleroderma; dermatomyositis; pemphigus.  
XX  
OS Arachis hypogaea.  
XX  
PN US2003235594-A1.  
XX  
PD 25-DEC-2003.  
XX  
PF 17-SEP-2002; 2002US-00245871.  
XX  
PR 14-SEP-1999; 99US-00396813.  
PR 17-JUL-2002; 2002US-00197000.  
XX  
PA (ANTI-) ANTIGEN EXPRESS INC.  
XX  
PI Humphreys R, Xu M;  
XX  
PS WPI; 2004-070554/07.  
XX  
DR Novel II-Key/antigen presentation enhancing hybrid polypeptide, useful  
PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus  
PT erythematosus and diabetes mellitus.  
XX  
PS Example 3; Page 21; 87pp; English.  
XX  
CC The invention relates to a novel antigen presentation enhancing hybrid  
CC polypeptide. The novel polypeptide has an N-terminal element consisting  
CC of 4-16 residues of a mammalian II-Key peptide and its non-N-terminal  
CC deletion modifications, a chemical structure covalently linking the N-  
CC terminal element to an MHC class II-presented epitope of a C-terminal  
CC element. The C-terminal element comprises an antigenic epitope, which  
CC binds to an antigenic peptide binding site of an MHC class II molecule.  
CC The antigen presentation enhancing hybrid polypeptide has the following  
CC activities: antibacterial, virucide, fungicide, antirheumatic,  
CC antiarthritic, neuroprotective, dermatological, immunosuppressive,  
CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation  
CC enhancing hybrid polypeptide is useful for modulating the immune response  
CC in an individual and for treating infections (such as bacteria, virus,  
CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus

CC erythematous, diabetes mellitus, myasthenia gravis, autoimmune  
 CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence  
 CC represents a mammalian II key related protein of the invention.

XX  
 SQ Sequence 507 AA;

Alignment Scores:  
 Pred. No.: 2,046-235 Length: 507  
 Score: 2665.00 Matches: 506  
 Percent Similarity: 99.80% Conservatives: 0  
 Best Local Similarity: 99.80% Mismatches: 1  
 Query Match: 99.03% Indels: 0  
 DB: 8 Gaps: 0

US-10-728-051-3 (1-1524) x ADM12139 (1-507)

QY	1	CGGACACACCGGAGAGACCGTCCAGTTCAGCGCTCAATCGCGAGACCTGAC	60
DB	1	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	20
QY	61	AATCGCATTTGAATCAGAGGCGGCTTACATTGAGACTTGGAAACCCCAACACGAGGATTC	120
DB	21	AsnArgIleGluSerGluGlyTyrlleGluThrTriAsnProAsnAsnGlnGluPhe	40
QY	121	GAATGCGCGGCTCGGCTCTCTCGCTTAGTCTCTCGCGCGCAACGCGCTTCGTAGGCT	180
DB	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro	60
QY	181	TTCTACTCCATGCTCCGAGGAGACTTCATCCAGCAGGAGGGATCTTGGTGG	240
DB	61	PheTySerAsnAlaProGlnGluIlePheIleGlnGlyArgGlyTyrlleGlyLeu	80
QY	241	ATATTCTCTGTTGCTTCTAGACTATGAGAGCTCACACAAAGTCTGCTCATCTCAG	300
DB	81	IlePheProGlyCysProArgHisTyrlleGluProHisThrGlnGlyArgSerGln	100
QY	301	TCCAAAGACCAACAGACTCTCCAGGAGAGAGCAAGCAACAGCAACGAGATGT	360
DB	101	SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	120
QY	361	CACGAGAGTGCACCGTTTCGATGAGGTGATCTCATTCAGTCTCCACCGTGTGCT	420
DB	121	HisGlnLysValHisArgThrAspGluGlyAspLeuIleAlaValProThrGlyValAla	140
QY	421	TTCTGCTCTACACCAACGACACTGATGTTGTTGCTGTTTCTCTACTGACCAAC	480
DB	141	PheTriLeuTyrlleAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn	160
QY	481	AACACGACCAACGAGTTGATCAGTTCCCGAGGAGATTCAATTTGGCTGGGAAACAGGAG	540
DB	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu	180
QY	541	CAAGAGTTCTTAAGGTACACGACCAACAGCAGCAAGCAGACGAGAGCTTACCATAT	600
DB	181	GlnGluPheLeuArgTyrlleGlnGlnSerArgGlnSerArgArgSerLeuProTyrlle	200
QY	601	AGCCCATACACCGCGCAAGTTCAGCTCAGACAGAGAGCGTGAATTTAGCCCTCGAGGA	660
DB	201	SerProTyrlleSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	220
QY	661	CAGCAGCGCGAGAACGAGAGGAGCAAGAGAGAAACGAGTGGAAACATCTTC	720
DB	221	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe	240
QY	721	AGCGGTTTCAGCGGAGTTCCTGGACACAGCTTCAGGTTTCAGCAGACAGAGATGTG	780
DB	241	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal	260
QY	781	CAAAACCTTAAGAGCGAGACCGAGAGTGAAGAGAGGAGGCGCATTTGTGACAGTGAGGGA	840
DB	261	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	280
QY	841	GGCCTCAGATCTTTGAGCCCCAGATAGAAAGAGAGCTGCCAGCAGAGAGAGATACGAT	900

DB	281	GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGlyArgAsp	300
QY	901	GAAAGATGAATATGAATACGATCAAGAGGATGAAGCGTGGCAGGAGGAGCAGAGCAGG	960
DB	301	GluAspGluTyrlleGluThrAspGluAspArgArgGlyArgGlySerArgGlyArg	320
QY	961	GGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAAAGAACATTTGGTGAAC	1020
DB	321	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	340
QY	1021	AGATCCCTGACATCTACACCTCAAGCTGTTCTCACTCAAACTGCCAACCATCTCAAC	1080
DB	341	ArgSerProAspIleTyrlleAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	360
QY	1081	CTTCTAATACTTAGTGGCTTGGACCTAGTGTGTAATATGAAATCTCTACAGGAATGCA	1140
DB	361	LeuLeuIleLeuArgTyrlleGlyProSerAlaGluTyrlleGlyAsnLeuTyrlleArgAsnAla	380
QY	1141	TTGTTTGTCTCCTCATTACCAACCAACGACACAGCATCATATATCGATTGAGGGACGG	1200
DB	381	LeuPheValAlaHisTyrlleAsnThrAsnAlaHisSerIleIleTyrlleArgLeuArgGlyArg	400
QY	1201	GCTCAGTCAAGTCTGTCAGACGACGACGACGAGTGTACGAGGAGCTTCAAGAG	1260
DB	401	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrlleAspGluGluGlnGlu	420
QY	1261	GGTCACTGCTTGTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCCGAGGAGCAGAAC	1320
DB	421	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	440
QY	1321	TTTCGAATACGTGGCATTTCAAGACAGACTCAAGGCGGAGATAGCCCACTCGCGGTGAA	1380
DB	441	PheGluTyrlleValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	460
QY	1381	AACTCCGTATAGATAACCTCGCGGAGGAGTGTTCGAAATTCATATGCGCTCCAAAGG	1440
DB	461	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrlleGlyLeuGlnArg	480
QY	1441	GAGCAGCAGGAGCAGCTTAAGAACACAAACCCCTTCAAGTCTTCTGTTCCACCGTCTCAG	1500
DB	481	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln	500
QY	1501	CAGTCTCCGAGGCTGTGGCT	1521
DB	501	GlnSerProArgAlaValAla	507

RESULT 5  
 AAY15246  
 ID AAY15246 standard; protein; 510 AA.  
 XX  
 AC AAY15246;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 09-NOV-1999 (first entry)  
 XX  
 DE Peanut allergen, Ara h 3, amino acid sequence.  
 XX  
 KW allergy; immune response; transgenic; allergen; epitope;  
 KW immunoglobulin E; Ig E; binding site; peanut.  
 XX  
 OS Arachis hypogaea.  
 XX  
 PN WO9938978-A1.  
 XX  
 PD 05-AUG-1999.  
 XX  
 PF 29-JAN-1999; 99WO-US002031.  
 PR 31-JAN-1998; 98US-0073283P.  
 PR 13-FEB-1998; 98US-0074590P.  
 PR 13-FEB-1998; 98US-0074624P.  
 PR 13-FEB-1998; 98US-0074633P.

```
PR 27-AUG-1998; 98US-00141220.
XX (UYAR-) UNIV ARKANSAS.
PA (UYNV) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
PA (SOSI/) SOSIN H.
XX Sosin H, Bannon GA, Burks AW, Sampson HA;
XX WPI; 1999-479189/40.
DR N-PSDB; AAZ06384.
XX Modified allergen with reduced IgE binding, useful for treating e.g.
XX allergies.
XX Disclosure; Page 39-40; 46pp; English.
XX This is the amino acid sequence of the Ara h 3 protein from Arachis
CC hypogea. The Ara h 3 protein has 4 IgE (Immunoglobulin E) binding
CC epitopes, one of which is immunodominant (AAV15281). Note: The first
CC three amino acids of this sequence are not encoded by the related
CC nucleotide sequence. By modifying the IgE binding sites the ability
CC of the allergen to provoke an immune response is downregulated. The epitopes
CC of the IgE binding sites can therefore be modified in genetically
CC engineered plants and animals to elicit less of an allergic response.
CC (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 510 AA;
Alignment Scores:
Pred. No.: 3,13e-235 Length: 510
Score: 2663.00 Matches: 506
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 98.96% Indels: 0
DB: 2 Gaps: 0
US-10-728-051-3 (1-1524) x AAY15246 (1-510)
QY 1 CGGCAGCAACGGAGAGAACGGCTGCCAGTTCAGCGCCTCAATGCGCAGACCTGAC 60
DB 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
QY 61 AATCGCATTGATCAGAGGGGGTTACATTGAGACTTGGAACTGGACCCCAACACAGGAGTTC 120
DB 24 AsnArgileGluSerGluGlyGlyTyriLeGluThrTrpAsnProAsnAsnGlnGluPhe 43
QY 121 GAATGCGCGCGTCGCTCTCTCGCTTAGTCTCCGCGCAACGCGCTTCGTAGGCGCT 180
DB 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 63
QY 181 TTCTACTCCAATGTCCTCCAGGAGATCTTCATCCAGCAAGAGGGGATCTTTGGGTTG 240
DB 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnArgGlyTyrPheGlyLeu 83
QY 241 ATATTCCTGGTGTCTCTAGACATATCAAGAGCTCACACAGGTCTCGATCTCAG 300
DB 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlnArgSerGln 103
QY 301 TCCCAAGACACCAAGACGCTCCAAAGGAGAAGCAAAAGCCCAACAGCAACGAGATGT 360
DB 104 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
QY 361 CACCAAGAGGTGACCGCTTTTCGATGAGGGTATCTCATTCAGTTCACCGGTTGCT 420
DB 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY 421 TTCTGGCTCTACACGACGACGACACTGATGTTGCTGCTTCTTCTTACTGACACCAAC 480
DB 144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
QY 481 AACACACGACACGACCTTGATTCAGTTCCTCCAGGAGATTCATTTGGCTGGACACCGAG 540
DB 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
QY 541 CAAGAGTTCCTTAAGTTACAGCAACAAAGCAGACAAAGCAGACAAAGCAGACCTTACCATAT 600
DB 184 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
QY 601 AGCCCATACAGCCGCGAAAGTACAGCTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
DB 204 SerProTyrSerProGlnSerGlnProArgGlnGlnGluArgGluPheSerProArgGly 223
QY 661 CAGCAGAGCCGCGAGAACGAGCAGGACGAGCAAGAAAGAAACGAGGTGGAACATCTTC 720
DB 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGlyGlyAsnIlePhe 243
QY 721 AGCGGCTTCAGCGCGAGTTCTCTGGAACAAGCCCTTCAGGTTGACGACGACAGATAGTG 780
DB 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
QY 781 CAAAACTTAGAGCGCAGAGCCGAGAGTGAAGAGAGAGAGGCCCATTTGTGACAGTGAGGGA 840
DB 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 283
QY 841 GGCTCTCAGAAATCTTGAGCCCGAGATAGAAAGAGACGTCGCCGACGAAAGAGGAATACGAT 900
DB 284 GlyLeuArgileLeuSerProAspArgLysArgAlaAspGluGluGluGluTyrAsp 303
QY 901 GAAGATGATATGATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 304 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 323
QY 961 GGGATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343
QY 1021 AGATCCCTCAGCATCTACAACTCTCAAGCTGGTTCTACTCAAACTGCAACAGATCTCAAC 1080
DB 344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
QY 1081 CTTCTAATATCTTATGCTGCTTGGACCTAGTGTGTAATATGTAATATCTCTCAGGAATGCA 1140
DB 364 LeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383
QY 1141 TTGTTTCTCCTCCTACACACCAACGACGACGACGACGACGACGACGACGACGACGACGACG 1200
DB 384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
QY 1201 GCTCAGCTGCAAGTCTGTCGAGCAGCAACGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
DB 404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423
QY 1261 GGTCACTGCTTGTGTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCCAGAGCGAGAAC 1320
DB 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
QY 1321 TTGCAATACCTGTCGCTTCAAGACGACTCAAGGCCCGCAGCATAGCCAACTCCCGGTGAA 1380
DB 444 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
QY 1381 AACTCCCTCATAGATAACCTGCGGAGGAGTGTGTCGAAATTCATATGCGCTCCAAAGG 1440
DB 464 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 483
QY 1441 GAGCAGCAAGGAGCTTAAGAACCAACACCCCTTCAAGTTCCTGTTCCACCCCTCTCAG 1500
DB 484 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 503
QY 1501 CAGTCTCCGAGGCGTGGCT 1521
DB 504 GlnSerProArgAlaValAla 510
RESULT 6
ID AAY40912
XX AAY40912 standard; protein; 512 AA.
```



Db 404 GlyArgAlaHisValGlnValValAspSerAsnGlyAsnArgValTyAspGluGluLeu 423  
QY 1255 CAAGAGGTCACGTGTTGGTGCCACAGAACTTCGCGCTCGCTGAAAGTCCACAGAGC 1314  
Db 424 GlnGluGlyHisValLeuValProGlnAsnPheAlaValAlaGlyLysSerGlnSer 443  
QY 1315 GAGAACTTCGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCACCTCGCC 1374  
Db 444 GluAsnPheGluTyValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAla 463  
QY 1375 GGTGAAAACTCCGTCATAGATAACCTCGCGAGAGGTGGTGCATAATTCATATGGCCTC 1434  
Db 464 GlyGluAsnSerValIleAspAsnLeuProGluValValAlaAsnSerTyGlyLeu 483  
QY 1435 CAAAGGAGCAGCAGCGCAGCTTAAGAAACAACACCCCTCAAGTTCTTCGTTCACCG 1494  
Db 484 GlnArgGluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProPro 503  
QY 1495 TCTCAGCAGTCTCGAGGGCTGGCT 1521  
Db 504 SerGlnGlnSerProArgAlaValAla 512

RESULT 7  
AAB33601

ID AAB33601 standard; protein; 526 AA.

XX AAB33601;  
AC  
XX  
XX 12-SEP-2003 (revised)  
DT 22-JAN-2001 (first entry)  
XX  
XX Modified Ara h 3 amino acid sequence.  
DE  
XX  
XX Allergy; allergic reaction; allergen; anaphylactic antigen; peanut;  
KW Ara h 1; Ara h 2; Ara h 3; epitope; binding; immunoglobulin E; IgE;  
KW food antigen; sensitising; immune response; anti-allergic.  
XX

OS Arachis hypogaea.

XX  
XX WO200051647-A2.  
PN  
XX  
XX 08-SEP-2000.PD

XX  
XX 03-MAR-2000; 2000WO-US005655.  
PF  
XX  
XX 03-MAR-1999; 99US-0122960P.  
PR  
XX  
XX 06-DEC-1999; 99US-00455294.  
PR

XX  
XX (MOUN ) MOUNT SINAI SCHOOL MEDICINE.  
PA  
XX  
XX Sampson HA;  
PI  
XX  
XX WPI; 2000-611341/58.  
DR

XX  
XX Non-human animal sensitized to an antigen, useful as an animal model for  
PT studying allergic reactions to allergens, such as those in food and in  
PT the environment.  
PT  
XX  
XX Example 6; Fig 17C; 124pp; English.

XX  
XX The present invention describes an animal model which can be used for  
CC studying allergic reactions to allergens. The animal is sensitised to a  
CC selected antigen by administering the antigen itself or a nucleic acid  
CC encoding the antigen, where preferably the antigen is an anaphylactic  
CC antigen. The sensitised animal can then be used to screen for compounds  
CC which help to prevent, ameliorate, or cure allergic conditions in  
CC humans. The animal model can be used for studying allergic reactions to  
CC allergens, such as those in food (peanuts, fruit, berry, nut, bean, milk,  
CC dairy products), or in the environment (weed pollen, grass pollen, tree  
CC pollen, mite, animal, animal dander, fungal, and insect antigens).

XX  
XX AAB33478 to AAB33601 represent sequences which are used in examples from  
CC the present invention to specifically examine the peanut allergy, and the

CC peanut anaphylactic antigens Ara h 1 to 3. (Updated on 12-SEP-2003 to  
CC standardise OS field)  
XX  
SQ Sequence 526 AA;  
Alignment Scores:  
Pred. No.: 6,27e-233 Length: 526  
Score: 2638.00 Matches: 502  
Percent Similarity: 99.01% Conservative: 0  
Best Local Similarity: 99.01% Mismatches: 5  
Query Match: 98.03% Indels: 0  
DB: 3 Gaps: 0

US-10-728-051-3 (1-1524) x AAB33601 (1-526)

QY 1 CGCAGCAACCGGAGGAGAAACCGTCCAGTTCACGCGCTCAATGCGCAGACCTGAC 60  
Db 5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24  
QY 61 AATCGCATTCGAATCAGAGGGCGGTACATTGAGACTTGGACCCCAACACACGAGGATTC 120  
Db 25 AsnArgIleGluSerGluGlyGlyTyIleGluThrTrpAsnAlaAsnGlnGluPhe 44  
QY 121 GAATGCGCGGCGTCCGCTCTCTCGTTAGTCTCTCGCGCAACGCGCTTCGTAGGCCT 180  
Db 45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgPro 64  
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGGGATCTTTGGGTG 240  
Db 65 PheTy-SerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyPheGlyLeu 84  
QY 241 ATATTCCCTGGTGTCTAGACACTATGAGAGCCTCACACAGGTCGTGATCTCTCAG 300  
Db 85 IlePheProGlyCysProArgHisTyGluGluProHisThrGlnGlyArgArgSerGln 104  
QY 301 TCCCAAGACACCAAGACGTCTCCAGGAGAGAGACCAAGCAACAGCAACAGCAGATAGT 360  
Db 105 SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124  
QY 361 CACCAGAGGTGCACCGTTTCGATAGGGGTATCTTCATTCAGTTCCTCCCGGTGTGCT 420  
Db 125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144  
QY 421 TTCTGCTCTACCAACGACACGACACTGATGTTGTGCTGTTCTTCTTACTGACACCAAC 480  
Db 145 PheTrpLeuTyArgAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164  
QY 481 AACACGACAAACCGCTTGATCAGTTCCCGCAGGAGATTCAATTTGGCTGGGAAACACGAG 540  
Db 165 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 184  
QY 541 CAAAGAGTCTTAAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600  
Db 185 GlnGluPheLeuArgTyGlnGlnSerArgGlnSerArgArgArgSerLeuProTy 204  
QY 601 AGCCCATACAGCCCGCAAGTACGCTAGACAGAGAGAGCGTGAAATTTAGCCCTCAGGA 660  
Db 205 SerProTySerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 224  
QY 661 CAGCAGCGCGCAGAGAACGAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
Db 225 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGlnGluGlnGlyAsnIlePhe 244  
QY 721 AGCGGCTTCAAGCGCGGAGTTCTTCGGAACAAGCCCTTCAGGTTGACGACAGACAGATAGT 780  
Db 245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspArgGlnIleVal 264  
QY 781 CAAAACCTTAAGAGCGGAGACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
Db 265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 284  
QY 841 GGCTCAGATCTTGAGCCCGCAGATAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 900

Db 285 GlyLeuArgAlaLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp 304  
QY 901 GAAGATCAATATCAATACGATGAAGAGGCTGCAGGCGAAGCAGAGGCGAG 960  
Db 305 GluAspGluTyrAlaTyrAspGluGluAspArgArgGlySerArgGlyArg 324  
QY 961 GGAAATGGTATTGAAGAGAGCATCTGCACCGCAAGTGCTAAAAAGAACATTTGTAGAAAC 1020  
Db 325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 344  
QY 1021 AGATCCCTCAGATCTACACCTCTCAGCTGGTTCATCTCAAACTGCCAACGATCTCAAC 1080  
Db 345 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364  
QY 1081 CTTCTAATACTAGGTGGCTTGGACCTAGTGTCTGAATATGAAATCTCTACAGCAATGCA 1140  
Db 365 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 384  
QY 1141 TTGTTTGTCTCCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200  
Db 385 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 404  
QY 1201 GCTCAGTGCAGCTCTGGACACCAACGCGCAACAGAGTGTACACGAGGAGCTTCAAGAG 1260  
Db 405 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 424  
QY 1261 GGTCACTGCTGTGTGTGTCACAGAACTTCGCGCTGCTGGAAGTCCACAGACGAGAAC 1320  
Db 425 GlyHisValLeuValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444  
QY 1321 TTGCAATACGTGCAATCAAGACAGACTCAAGCCCGACGATACCCAACTTCGCGGTGAA 1380  
Db 445 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464  
QY 1381 AACTCCGTCTAGATACCTCGCGGAGGAGTGTTCGCAATTCATATGCTCCAAAGG 1440  
Db 465 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 484  
QY 1441 GAGCAGGCAAGGAGCTTAAAGAACAAACACCCCTTCAAGTTCCTTCACCGTCTCTAG 1500  
Db 485 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln 504  
QY 1501 CAGTCTCCGAGGGCTGTGGCT 1521  
Db 505 GlnSerProArgAlaValAla 511

RESULT 8  
AAU04711  
ID AAU04711 standard; protein; 526 AA.  
XX  
AC AAU04711;  
XX  
XX  
DT 23-OCT-2001 (first entry)  
XX  
DE Modified anaphylactic antigen Ara h 3.  
XX  
XX Ara h 3; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;  
KW allergy; mast cell; basophil; mouse.  
XX  
XX Mus sp.  
OS Synthetic.  
XX  
XX W0200140264-A2.  
XX  
XX 07-JUN-2001.  
XX  
XX 06-DEC-2000; 2000WO-US033124.  
XX  
XX 06-DEC-1999; 99US-00455294.  
PR 23-JUN-2000; 2000US-0213765P.  
PR 27-SEP-2000; 2000US-0235797P.  
XX  
XX (PANA-) PANACEA PHARM LLC.

(UYAR-) UNIV ARKANSAS.  
(MOUN ) MOUNT SINAI SCHOOL MEDICINE NEW YORK.

Bannon CA, Burks WA, Caplan MJ, Sampson H, Sosin H;  
WPI; 2001-381378/40.

Antigenic fragments useful for reducing anaphylactic risk and reducing  
the severity and/or number of allergic symptoms in individuals sensitive  
to antigens, have reduced ability to bind immunoglobulin E.

Disclosure; Fig 12; 100pp; English.

The sequence represents the amino acid sequence of modified anaphylactic  
antigen Ara h 3, which has been altered to disrupt immunoglobulin E (IgE)  
binding sites. Ara h 3 is an anaphylactic antigen (A), which was used to  
design antigenic peptides having a reduced ability to bind IgE as  
compared with the intact (A), or having a sequence substantially  
identical to a portion of sequence of an antigen that includes at least  
one IgE binding site, where at least one IgE binding site of the peptide  
is altered. The antigenic peptides are used in a composition which is  
useful for reducing risk or severity of allergic reaction to an antigen.  
This is done by identifying an individual at risk of allergic reaction to  
an antigen by identifying prior display of allergic symptoms when exposed  
to the antigen, or a familial relationship with an individual who  
previously displayed allergic symptoms when exposed to the antigen.  
Following this an antigen-specific IgE present on one or more mast cells  
or basophils in the individual's serum is identified. The individual is  
then contacted with a peptide corresponding to a portion of the antigen,  
which is selected, formulated, and delivered so that binding of the  
peptide to antigen-specific IgE is reduced as compared with IgE binding  
of intact antigen. The composition is also useful for treating and  
preventing allergic reactions

SQ Sequence 526 AA;

#### Alignment Scores:

Pred. No.:	6,27e-233	Length:	526
Score:	2638.00	Matches:	502
Percent Similarity:	99.01%	Conservative:	0
Best Local Similarity:	99.01%	Mismatches:	5
Query Match:	98.03%	Indels:	0
DB:	4	Gaps:	0

US-10-728-051-3 (1-1524) x AAU04711 (1-526)

QY	1	CGGCAGACACCGAGGAGAACCGTGCAGGCTCCAGCGCTCAATCGCAGAGACTGAC	60
Db	5	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	24
QY	61	AATCGCATTGAATCAGAGGGCGGTTACATTGAGACTTGGAAACCCCAACACGAGGTTTC	120
Db	25	AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnAlaAsnGlnGluPhe	44
QY	121	GAATGCGCGGCTGCGCTCTCTCTGCTTAGTCTTCGCGCGCAAGCCCTTCGTAGGCTT	180
Db	45	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	64
QY	181	TTCTACTCCCAATGCTCCCGAGGAGATCTTCATCCAGAGGAGGAGTACTTTGGGTG	240
Db	65	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	84
QY	241	ATATTCCCTGGTGTCTTAGACACTATGAAGAGCTTCACACAAAGGTCTCGATCTCAG	300
Db	85	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln	104
QY	301	TCCAAAGACCAACAGACCTCTCCAGAGAGAACCAAGCCACAGCAACAGTAGT	360
Db	105	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	124
QY	361	CACCAAGAGTGCACCGTTTCGATGAGGTGATCTCATTGCAGTTCACACCGGTGTGCT	420
Db	125	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	144







US-10-728-051-3 (1-1524) x AAU05036 (1-526)	
QY	1 CGGACCAACCGAGAGAACCGGTGCCAGTTCCAGCGCCTCAATGCCGAGAGACCTGCAC 60
Db	5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24
QY	61 AATCGCATTGAATCAGAGCGCGTTCATTAGACTTGGAAACCCCAACACAGAGATTC 120
Db	25 AsnArgIleIleuSerGluGlyTyrIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44
QY	121 GAATGCGCGGCGTCCCTCTCTCGCTTAGTCTCCCGCCGCAAGCCCTTCGTAGGCT 180
Db	45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 64
QY	181 TTCTACTCCATGCTCCCGAGGATCTTCATCCAGCAAGGAGGGATACCTTGGGTG 240
Db	65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 84
QY	241 ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCCTCACACAAAGCTCGTCTCAG 300
Db	85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 104
QY	301 TCCCAAGACCCACCAAGAGCTTCCAGGAGAAAGACCAAGCCCAACAGCAACGAGATAGT 360
Db	105 SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124
QY	361 CACAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
Db	125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144
QY	421 TTCTGGCTCTACCAACCAACGACACTGATGTTGTTGTTGTTCTCTTACTGACCAAC 480
Db	145 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164
QY	481 AACAAACCAACGAGTTGATGTTCCCGAGAGATTCATTTGCTGGGAACAGGAG 540
Db	165 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 184
QY	541 CAAGAGTTCTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
Db	185 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr 204
QY	601 AGCCCATACCGCCGAAAGTACGCTTAGACAGAGCGTGAATTTAGCCCTCAGGGA 660
Db	205 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 224
QY	661 CAGCACAGCCGACAGAACGACGACGAGCAAGAAACGAAAGGTGAAACATCTTC 720
Db	225 GlnHisSerArgArgGluAArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 244
QY	721 AGCGGCTTCACGCGGAGTTCTCTGAAACAAAGCCTTCCAGGTTACGACAGACAGATAGT 780
Db	245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal 264
QY	781 CAAACCTTAAGACGCGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAGGGA 840
Db	265 GlnAsnLeuArgGlyGluThrGluSerGluGluGlyAlaIleValThrValArgGly 284
QY	841 GGCCTCAGATCTTTGACCCAGATAGAAAGAGAGCTGCCGACGAAGAGGAATACGAT 900
Db	285 GlyLeuArgAlaLeuSerProAspArgLysArgArgAlaAspGluGluGluGluTyrAsp 304
QY	901 CAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGAGGAGG 960
Db	305 GluAspGluTyrAlaTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 324
QY	961 GGGATGTGATTAAGACAGACGATCTGCACCCCAAGTCTTAAAGAAACATTTGTTAGAAC 1020
Db	325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 344
QY	1021 AGATCCCTGACATCTCAACCTCAAGCTGGTTCACTCAAACTGCCCAACGATCTCAAC 1080
Db	345 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364
RESULT 10	
ID	AD038357
XX	AD038357 standard; protein; 507 AA.
AC	AD038357;
DT	15-JUL-2004 (first entry)
XX	Peanut allergen Ara h 3.
DE	
XX	Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;
KW	Antirheumatic; Neuroprotective; Antiinflammatory; Dermatological;
KW	Immunosuppressive; Antidiabetic; Antithyroid; Antisthmatic;
KW	Antiallergic; Cytostatic; Antipsoriatic; Gene Therapy; Vaccine;
KW	MHC Class II; Ii-key motif; immune response; anthrax; EBOLA; HIV;
KW	influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;
KW	rickettsia; rheumatoid arthritis; multiple sclerosis;
KW	lupus erythematosus; diabetes mellitus; myasthenia gravis;
KW	autoimmune thyroiditis; scleroderma; dermatomyositis; pemphigus; asthma;
KW	allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;
XX	adenoma; peanut; Ara h 3.
XX	Arachis hypogaea.
OS	
XX	US2004058881-A1.
PN	
XX	25-MAR-2004.
PD	
XX	
PF	24-SEP-2002; 2002US-00253286.
XX	
PR	24-SEP-2002; 2002US-00253286.
XX	
XX	(ANTI-) ANTIGEN EXPRESS INC.
PA	
XX	
PI	Humphreys RE, Xu M;
XX	
DR	WPI; 2004-294259/27.
XX	
PT	New non-naturally occurring protein or polypeptide modified by recombinant DNA techniques, useful for treating multiple sclerosis,

diabetes mellitus, myasthenia gravis, scleroderma, allergic rhinitis, colitis, cancer or psoriasis.

Example 3; Page 21-22; 90pp; English.

The invention relates to a non-naturally occurring protein or polypeptide (I) modified by recombinant DNA techniques comprising: a C-terminal element comprising an MHC Class II-presented epitope; an N-terminal element comprising an Ii-key motif; and an intervening element comprising a sequence of 4-11 amino acid residues where the modification by recombinant DNA techniques taking place within elements (b) and (c). Also described are methods for: suppressing or enhancing an immune response directed toward an MHC (major histocompatibility complex) Class II-presented epitope of interest. Suppressing an immune response directed toward an MHC Class II-presented epitope of interest comprises: providing a nucleic acid sequence encoding the MHC Class II-presented epitope of interest, the nucleic acid sequence encoding an Ii-key motif located 4-11 amino acids upstream from the N-terminal residue of the MHC Class II-presented epitope of interest; and modifying the Ii-key motif to decrease its conformance to the archetypal Ii-key regulatory motif. Enhancing an immune response directed toward an MHC Class II-presented epitope of interest comprises: providing a nucleic acid sequence encoding the MHC Class II-presented epitope of interest, the nucleic acid sequence lacking an Ii-key motif located 4-11 amino acids upstream from the N-terminal residue of the MHC Class II-presented epitope of interest; and modifying the Ii-key motif to increase its conformance to the archetypal Ii-key regulatory motif. The protein or polypeptide of interest corresponds to a protein or polypeptide encoded by an infectious pathogen selected from anthrax, EBOLA, HIV or influenza, preferably vaccinia virus. The non-naturally occurring protein or polypeptide (I) modified by recombinant DNA techniques is useful for treating infectious diseases caused or associated with infection by a bacterium, virus, parasite, fungus, rickettsia or other infectious agents. It is also useful for treating rheumatoid arthritis, multiple sclerosis, lupus erythematosus, diabetes mellitus, myasthenia gravis, autoimmune thyroiditis, scleroderma, dermatomyositis, pemphigus, asthma, allergic rhinitis, topical dermatitis, colitis, cancer, psoriasis or adenomas. The present sequence represents the amino acid sequence of peanut allergen Ara h 3 used in the invention.

Sequence 507 AA:

**Alignment Scores:**

Pred. No.:	8e-231	Length:	507
Score:	2615.00	Matches:	498
Percent Similarity:	98.22%	Conservative:	0
Best Local Similarity:	98.22%	Mismatches:	9
Query Match:	97.18%	Indels:	0
DB:	8	Gaps:	0

US-10-728-051-3 (1-1524) x ADO38357 (1-507)

1	CGGCAGCAACCGGAGGAGAA	CGCGTGCAGGTTCCAGCGCTCAATGCGCAGAGACCTGAC	60
1	ArgGlnGlnProGluGluAsnAla	CysGlnPheGlyArgLeuAsnAlaGlnArgProAsp	20
61	AATCGCAATTGAATCAGAGGGCGGTTACACTTGAGACTTGGAACCCCAACCAACAGAGAGTTC	120	
21	AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlyGluPhe	40	
121	GAATGCGCGCGGCTGCGCCTCTCTCGCTTAGTCTCTCGCGCGGAACGCCCTTCGTAGGCGCT	180	
41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	60	
181	TTCTACTCCAAATGCTCCCGCAGAGACTTTCATCCAGCAAGGAAGGGGATACTTTCGGGTG	240	
61	PheTySerAsnAlaProGlnGluIlePheIleGlyGlyArgGlyTyrlPheGlyLeu	80	
241	ATATTCCCTGGTTGCTTAGACACTATGAAGAGCCTCACACAAGAGTCGTGCATCTCAG	300	
81	IlePheProGlyCysProArgHisTyrlGluGluProHisThrGlyGlyArgArgSerGln	100	
301	TCCCAAGACCAACCAAGAGCTCTCCAGGAGGAAGACCAAGGCCAACGACGAGTACT	360	

```

Db      461 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 480
QY      1441 GAGCAGCGCAGGAGCTTAAAGAACACACCCCTTCAAGTTCTTCGTTCCACCGTCTTCAG 1500
Db      481 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPheValProProSerGln 500
QY      1501 CAGTCTCCGAGGGCTGTGGCT 1521
Db      501 GlnSerProArgAlaValAla 507

RESULT 11
ABG71266
ID      ABG71266 standard; protein; 481 AA.
XX
AC      ABG71266;
DT      17-DEC-2002 (first entry)
XX
XX      Glycine max (Soybean) var. Dare protein.
DE      Soybean; Glycinin; atomic coordinate data; processability; soya protein;
KW      Dare; protein co-ordinate data.
XX
XX      Glycine max.
OS
XX      JP2002193996-A.
XX
XX      10-JUL-2002.
XX
XX      21-DEC-2000; 2000JP-00405097.
XX
XX      21-DEC-2000; 2000JP-00405097.
XX
XX      (KYOU ) UNIV KYOTO.
XX
XX      WPI; 2002-685438/74.
DR      N-PSDB; ABS55193.
XX
XX      Glycinin, beta-conglycinin and proglycinin, their crystal structures,
PT      three dimensional coordinates, three dimensional structured and models
PT      and their uses.
XX
XX      Disclosure; Page 1273-1274; 1298pp; Japanese.
XX
XX      The present invention relates to a new Glycinin characterised by the
CC      atomic coordinate data fully defined in the specification. The structure
CC      can be used for improving processability of soya protein. The present
CC      amino acid sequence represents the Glycine max (Soybean) var. Dare
CC      protein, as described in the specification
XX
SQ      Sequence 481 AA;

Alignment Scores:
Pred. No.:      1 32e-128      Length:      481
Score:          1502.00      Matches:      297
Percent Similarity: 71.57%      Conservative: 68
Best Local Similarity: 58.24%      Mismatches: 91
Query Match:    55.82%      Indels:      54
DB:            5      Gaps:      7

US-10-728-051-3 (1-1524) x ABG71266 (1-481)

QY      1 CGGCAGCAACCGGAGGAGACGGGTCCAGTTCAGCGCTCAATGCGCAGACCTGAC 60
Db      23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
QY      61 AATCGCATTAATCAGAGGCGGTACATTGAGACTTGGAACTTGGAAACCCCAACACGAGGATTC 120
Db      43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY      121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCT 180
Db      63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82

```

```

QY      181 TTCTACTCAATGCTCTCCCGAGGAGATCTTTCATCCAGCAAGGAGGATACATTTGGGTTG 240
Db      83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
QY      241 ATATTCCCTGCTGCTAGACACTATGAAGAGCCCTCACACCAAGGTGTCGATCTCAG 300
Db      103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
QY      301 TCCAAAGACCAACCAAGACGCTCTCCAAGGAGAAGACCAAGCCCAACAGCAACGAGATAGT 360
Db      120 SerSerArgPro-----GlnAspArg 126
QY      361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTGCGAGTTCACCGGTGTGCT 420
Db      127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
QY      421 TTCTGGCTCTACAACGACCAACGACACTGATGTTGCTGTTCTCTTACTGACACCAAC 480
Db      147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
QY      481 AACCAACGACCAACGAGCTTCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAACACGGAG 540
Db      167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY      541 CAAGAGTTCTTAAGGTACCAACGACCAACGACCAACGACCAACGACCAAGAGCTTTACCATAT 600
Db      187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly-----199
QY      601 AGCCCATACAGCCCGCAAGTCCAGCTAGACCAAGAGAGCGGTGNAATTTAGCCCTCAGGGA 660
Db      200 -----Gly 200
QY      661 CAGCACAGCCGACAGAACGACGACGACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 720
Db      201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluGluGluGluGluGluGluGluGluGlu 220
QY      721 AGCGGCTTCACCGCGGAGTTCCTGGAAACAAAGCCTTCCAGGTTGACGACGACGACGATAGTG 780
Db      221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
QY      781 CAACACCTTAAGAGCGGACCGAGAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db      240 ArgLysLeuGlnGlyGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 259
QY      841 GGCTCAGAAATCTTACGACCA-----GATAGAAAGACGACGCTGCCGACCAAGAAAGAG 891
Db      260 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGluGluGluGlu 279
QY      892 GAATACGATGAAGATGAATATGAATACGATGAAGAGGATGAAGAGGCGTGGCAGGGGAAGC 951
Db      280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
QY      952 AGAGGCGGGGGAATGGTATTGAGAGACGATCTGCACCGCAAGTCTCTAAAGAAAGACATT 1011
Db      295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
QY      1012 GGTAGAAACAGATCCCTCAGCATCTACAACCTCTCAAGCTGGTTCACTCAAAACTGCCAAC 1071
Db      312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
QY      1072 GATCTCAACTTCTTAATATCTTAGTGGCTTGGACCTTAGTGTGAATATGGAATCTCTAC 1131
Db      332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
QY      1132 AGGAATGCAATTTGCTCGCTCACTACACCAACCAACGACGACGACGACGACGACGACGACGAC 1191
Db      352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
QY      1192 AGGGGACGGGCTCACCTGCAAGTCTGGGACAGCAACGGCAACAGAGAGTGTACGACGAGGAG 1251
Db      372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391

```



```
QY 541 CAAGAGTCTTTAAGTACCAAGCAAAAGCAGACGAAGAAGCTTTACCATAT 600
Db |||||
QY 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly-----199
Db |||||
QY 601 AGCCCATACAGCCCGCAAGTAGCTAGCAAGAAGAGCGTGAATTTAGCCCTCGAGGA 660
Db |||||
QY 200 -----Gly 200
Db |||||
QY 661 CAGCAGAGCGGCAGAGACGAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db |||||
QY 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
QY 721 AGCGGCTTACGCGCGGAGTTCTCGGAACAGCCCTTCAGGTTGACACACAGACAGATAGTG 780
Db |||||
QY 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
Db |||||
QY 781 CAAAACCTTAAGCGCAGAGCCGAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db |||||
QY 240 ArgLysLeuGlnGlyGluAsnGluGluGluGluLysGlyAlaIleValThrValLysGly 259
Db |||||
QY 841 GGCCTCAGATCTTGAGCCCA-----GATAGAAAGAGAGCTGCCGACGAAGAGAG 891
Db |||||
QY 260 GlyLeuSerValIleSerProProThrGluGluGlnGlnGlnArgProGluGluGluGlu 279
Db |||||
QY 892 GAATACGATGAAGATGAATATGAATACGATGAAGAGAGATAGAGAGCGCTGCGAGGGAAGC 951
Db |||||
QY 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
Db |||||
QY 952 AGAGCAGGGGGAATGGTATTGAAGAGAGAGATCTGCACGCAAGTCTTAAAGAACATT 1011
Db |||||
QY 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
Db |||||
QY 1012 GGTAGAAACAGATCCCTGACATCTACACCTCAAGCTGGTTCACCTCAAACTGCCAAC 1071
Db |||||
QY 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
Db |||||
QY 1072 GATCTCAACCTTCTAATCTAGTGGCTTGGACCTAGTCTGAATATGAATCTCTAC 1131
Db |||||
QY 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
Db |||||
QY 1132 AGAATGCAATGTTGTGCTCACTCAACACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1191
Db |||||
QY 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
Db |||||
QY 1192 AGGGGAGCGGCTCACTGCAAGTCGTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1251
Db |||||
QY 372 AsnGlyArgAlaLeuValGlnValAlaValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
Db |||||
QY 1252 CTTCAAGAGGGTCACTGCTGTGGTGGCCACAGAACTTCGCCCTCGCTCGAGAGTCCACAG 1311
Db |||||
QY 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
Db |||||
QY 1312 AGCAGAACTTGAATACGTGGCATTTCAAGCAGACTCAAGGCCCGAGCATAGCAACCTC 1371
Db |||||
QY 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
Db |||||
QY 1372 GCGGTGAAACCTCCCTCATAGATAACCTGCGGAGGAGTGTTGCCAATTCATATGCG 1431
Db |||||
QY 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
Db |||||
QY 1432 CTCCAAAGGAGCAGGAGCAGGAGCTTAAGAACCAACACCCCTTCAAGTTCTTCGTTCGA 1491
Db |||||
QY 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
Db |||||
QY 1492 CGTCTCAGAGCTTCGAGGCGCTGGCT 1521
Db |||||
QY 472 ProLysGluSerGlnArgArgValValAla 481
Db |||||
RESULT 13
ADL90187
ID ADL90187 standard; protein; 481 AA.
XX
```

```
AC ADL90187;
XX 20-MAY-2004 (first entry)
XX Soybean glycinin G3 protein.
XX immunomodulator; immunotherapy; allergen characterisation;
KW immunoglobulin E; allergen sensitivity; soybean; glycinin G3;
KW acidic protein.
XX Glycine max.
XX US2003166518-A1.
XX 04-SEP-2003.
XX 12-JAN-2001; 2001US-00759967.
XX 13-JAN-2000; 2000US-0175948P.
XX 03-MAR-2000; 2000US-0186724P.
XX (BEAR/) BEARDSLEE T A.
PA (ZEEC/) ZEECE M G.
PA (SARA/) SARATH G.
PA (MARK/) MARKWELL J P.
XX Beardslee TA, Zeece MG, Sarath G, Markwell JP;
XX WPI; 2003-898094/82.
XX Allergen characterization comprises obtaining a recombinant fusion
PT protein and detecting the binding of immunoglobulin E molecules in the
PT biological sample to the recombinant fusion protein.
XX Disclosure; SEQ ID NO 21; 34pp; English.
XX The invention describes a method of allergen characterisation comprising:
CC obtaining a recombinant fusion protein; attaching the recombinant fusion
CC protein to a substrate through the native protein; contacting the
CC recombinant fusion protein attached to the substrate with a biological
CC sample from an individual; and detecting the binding of immunoglobulin E
CC molecules in the biological sample to the recombinant fusion protein.
CC Also described are: a method for determining the sensitivity of an
CC individual to a suspected allergen; a method for determining the amount
CC of immunoglobulin E specific for an allergen in a biological sample; a
CC method of immunotherapy; a method of allergen characterisation; a method
CC for determining the sensitivity of an individual to a suspected allergen;
CC a method of determining the amount of immunoglobulin E specific for an
CC allergen in a biological sample; a kit comprising the recombinant fusion
CC protein and instructions for using the recombinant fusion protein to
CC determine IgE binding to the know or suspected allergen; and a method for
CC epitope determination. The method is useful for characterising allergens.
CC This is the amino acid sequence of soybean glycinin G2 acidic protein.
CC that can be used to demonstrate the methods of the invention.
XX Sequence 481 AA;
SQ
Alignment Scores:
Pred. No.: 1,32e-128 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.57% Conservative: 68
Best Local Similarity: 58.24% Mismatches: 91
Query Match: 55.82% Indels: 54
DB: 7 Gaps: 7
US-10-728-051-3 (1-1524) x ADL90187 (1-481)
QY 1 CGGCAGCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db |||||
QY 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
Db |||||
QY 61 AATCGCATTTGAATCAGAGCGCGGCTTACATTGAGACTTGGAAACCCCAACACGAGAGTTC 120
Db |||||
```

Db 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62  
Qy 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTATGCTCTCGCGCGCAACGCGCTTCGTAGGCCT 180  
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82  
Qy 181 TTCTACTCCATGCTCCCGAGGAGATCTTATCAGCAAGAGAGGAGGATCTTTGGGTTG 240  
Db 83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlySerGlyIlePheGlyMet 102  
Qy 241 ATATTCTCTGCTCTCTAGACACTATGAAGAGCTCACACACAGGTCTCGATCTCAG 300  
Db 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119  
Qy 301 TCCCAAGACACCAAGACGCTCTCAAGGAGAGACCAAGCAACACGAGATAGT 360  
Db 120 SerSerArgPro-----GlnAspArg 126  
Qy 361 CACAGAAGGTGACCGTTTCGATGAGGGTGAATCTCATTTGAGTTCCACCGGTGTGCT 420  
Db 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146  
Qy 421 TTCTGGCTTACACGACACGACACGATGATGCTGCTGCTTCTTACTTGACACCAAC 480  
Db 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166  
Qy 481 AACACACACACACGCTTGATCAGTCTCCCGAGGATTCATTTGGCTGGGAACACCGAG 540  
Db 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186  
Qy 541 CAGAGTTCTTAAGGTACAGCAACAAAGCAGACAAAGCAGACGAGCAAGAGCTTACCAT 600  
Db 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly----- 199  
Qy 601 AGCCATACACCGCGCAAGTACGCTAGACAGAGCGTGAATTTAGCCCTCGAGGA 660  
Db 200 -----Gly 200  
Qy 661 CAGCACAGCGCAGAACGACGACGAGCAAGAAAGAAACGAGGTGGAAACATCTTC 720  
Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGlnGluGlyGlySerIleLeu 220  
Qy 721 AGCGGCTTACGCGAGTTCTCTGGAAACAAGCTTCCAGGTTCAGCAGACAGATAGTG 780  
Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239  
Qy 781 CAAACCTAGACGCGGACCGCAGAGTGNAGAGGAGCGCATTTGACAGTTCAGGGGA 840  
Db 240 ArgLysLeuGlnGlyGluAsnGluGluGluGlyAlaIleValThrValLysGly 259  
Qy 841 GGCCTCAGATCTTTGAGCCCA-----GATAGAAAGAGACGTCGCGACGAGAGAG 891  
Db 260 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu 279  
Qy 892 GAATACGATGAAGATGAATATGAATACATGAAGAGATAGAGCGCTGGCAGGGGAAGC 951  
Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294  
Qy 952 AGAGCGAGGGGAATGTTATTAAGACAGACGATCTGACCGCAAGTGTCTAAAGAAACATT 1011  
Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311  
Qy 1012 GGTAGAAACAGATCCCTGACATCTACAACTCAAGCTGCTTCACTCAAACTGCCAAC 1071  
Db 312 GlyGlnThrSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 331  
Qy 1072 GATCTCAACCTTCTAATCTTAGTGTGGCTTGGACCTAGTGTGTAATATGGAATCTCTAC 1131  
Db 332 SerLeuAspPheProAlaLeuSerTyrLeuLysLeuSerAlaGlnPheGlySerLeuArg 351  
Qy 1132 AGGAATGATGTTTGTGCTGCTACATACAAACCAACGACGACATCATATATCGATTG 1191  
Db 352 LysAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371

Qy 1192 AGGGAGCGGCTCACGTGCAAGTCTGTGACAGCAACGCGCAACAGAGTGTACGACGAGG 1251  
Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391  
Qy 1252 CTTCAAGAGGTCACGTGCTTGGTGGTCCACACAACTTCGCGCTCGCTGGAAGTCCAG 1311  
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411  
Qy 1312 AGCGAGAACTTCGAATACGTGGCATTCAACAGACACTCAAGGCGCCAGCATAGCCAACTC 1371  
Db 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431  
Qy 1372 GCCGTTGAAACTTCGTCATAGATAACCTCCGAGGAGGTGTTGCAATTCATATGCGC 1431  
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451  
Qy 1432 CTCAAGAGGAGCAGCGCAGCTTAAGAACCAACCCCTCAAGTTCCTTCGTTCCA 1491  
Db 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471  
Qy 1492 CCGTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521  
Db 472 ProLysGluSerGlnArgArgValAla 481  
RESULT 14  
ADG43988  
ID ADG43988 standard; protein; 481 AA.  
XX AC ADG43988;  
XX DT 26-FEB-2004 (first entry)  
XX DE G. max glycinin G3 subunit protein.  
XX KW oil content; plant; storage protein; seed-specific promoter; 2S-albumin;  
KW 7S-globulin; 11S-globulin; 12S-globulin; zein-prolamine; transgenic;  
KW oil production; fat production; free fatty acid production; food;  
KW animal feed; pharmaceutical; fine chemical production; glycinin.  
XX OS Glycine max.  
XX PN WO2003077643-A2.  
XX PD 25-SEP-2003.  
XX PF 17-MAR-2003; 2003WO-EP002733.  
XX PR 20-MAR-2002; 2002DB-01012893.  
XX PS (BADI ) BASF PLANT SCI GMBH.  
XX PI Bauer J;  
XX DR WPI; 2004-011485/01.  
XX DR N-PSDB; ADG43987.  
XX PT Increasing total oil content of plants, useful e.g. as foods or animal  
PT feeds, by reducing amount of storage proteins, particularly with double-  
PT stranded interfering RNA.  
XX PS Claim 4; SEQ ID NO 28; 253pp; German.  
XX CC This invention describes a novel method for increasing the total oil  
CC content of a plant by reducing the amount of at least one storage protein  
CC in the plant (or its tissue, organs, parts or cells) and selecting plants  
CC that have higher total oil content than starting plants. The storage  
CC protein is suppressed by introducing antisense RNA, optionally combined  
CC with a ribozyme, sense RNA that induces co-suppression, DNA-binding  
CC factors directed against storage protein genes, viral sequences that  
CC degrade storage protein RNA, constructs that induce homologous  
CC recombination of endogenous storage protein genes or mutations into  
CC storage protein genes. Most preferably a plant cell is stably transfected



CC with a recombinant expression construct, then regenerated to plants that  
 CC express the incorporated sequence. The expression constructs particularly  
 CC contain a seed-specific promoter and they are introduced into plants by  
 CC standard methods, e.g. via Agrobacterium. The preferred storage proteins  
 CC of the invention are 2S-albumins, 7S or 11S/12S-globulins or zein-  
 CC prolamines. Transgenic organisms produced by the new method are used for  
 CC production of oils, fats, free fatty acids or their derivatives, useful  
 CC as foods, animal feeds, pharmaceuticals and fine chemicals. This sequence  
 CC represents a storage protein used to illustrate the method of the  
 CC invention.

XX  
 SQ Sequence 481 AA;

# Alignment Scores:

Pred. No.: 1,32e-128 Length: 481  
 Percent: 1502.00 Matches: 297  
 Score: 71.57% Conservative: 68  
 Best Local Similarity: 58.24% Mismatches: 91  
 Query Match: 55.82% Indels: 54  
 DB: 8 Gaps: 7

US-10-728-051-3 (1-1524) x ADG43988 (1-481)

QY	1	CGCAGACACCGGAGGAGACGGTCCAGTTCACGGCTCAATGGCGAGACCTGAC	60
DB	23	ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp	42
QY	61	AATCGCAATCAATCAGAGGCGGTTCATTCAGACTTGGAAACCCCAACACGAGAGTTC	120
DB	43	AsnArgIleGlnSerGluGlyGlyPheIleGluThrTrpAsnProAsnLysProPhe	62
QY	121	GAATGGCGCGGTGCGCTCTCTCGCTAGTCTCCGCGCGCAACGCGCTTCGTAGCCT	180
DB	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro	82
QY	181	TTCTACTCCAACTGCCAGGAGACTTCATCAGCAGAGGAGGAGTACTTTGGTTC	240
DB	83	SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet	102
QY	241	ATATTCCTCGTGTCTCTGACACTATGAGAGCGCTCACACAAAGTCTCGATCTCAG	300
DB	103	IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly	119
QY	301	TCCCAAGACCAACAGACGCTCCAGAGGAGAGACCAAGCCACAGCAACGAGATAGT	360
DB	120	SerSerArgPro	126
QY	361	CACCAGAGGTGCACCGTTCGATGAGGTGATCTCATTCAGTCCACCGGTGTGCT	420
DB	127	HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla	146
QY	421	TTCTGGCTCTACACGACGACACTGATGTGTGTGCTGTTCTCTTACTGACACCAAC	480
DB	147	TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn	166
QY	481	AACAAGACCAACGAGTGTGATGTTCCCGAGGAGATTCAATTTGGTGGGAACAGGAG	540
DB	167	SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	186
QY	541	CAAGAGTCTTAAGGTACACGACCAACAGCAGACCAAGCAGACGAGCAAGACCTTACCATAT	600
DB	187	GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly	199
QY	601	AGCCCATACAGCCCGCAAGTCCAGCTCAGACAGAGAGCGTGAATTTAGCCCTCGAGGA	660
DB	200	-----Gly 200	
QY	661	CAGCAGACCGCAGAGACGAGGAGGACAGAGAGAGAAACAGAGTGGAAACATCTTC	720
DB	201	ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGluAsnGluGlyGlySerIleLeu	220
QY	721	AGCGGCTTCACGCGGAGTTCCTGGAAACAGCCCTCCAGGTTCCAGCAGACAGATAGTG	780

DB	221	SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal	239
QY	781	CAAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGAGGGAGCCATTGTGACAGTGGGA	840
DB	240	ArgLysLeuGlnGlyLeuAsnGluGluGluLysGlyAlaIleValThrValLysGly	259
QY	841	GGCTCAGAACTTGGAGCCCA-----GATAGAAAGAGAGCTGCCACGAGAGAG	891
DB	260	GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu	279
QY	892	GAATACGATGAAGATGAATATGATGAGGATGAAGAGAGGCGTGCAGGGAAGC	951
DB	280	LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer	294
QY	952	AGAGCAGCGGGAATGTATTGAAGAGACGATCTGCACCGCAGTCTCTAAAGAACATT	1011
DB	295	Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle	311
QY	1012	GGTAGAAACAGATCCCTGACATCTACAACCTCAAGCTGGTTCACCTCAAACTGCCAAC	1071
DB	312	GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr	331
QY	1072	GATCTCAACCTTCTAATACTTAGTGGCTTGACCTAGTGTGCTGAATATGGAATCTCTAC	1131
DB	332	SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg	351
QY	1132	AGAAATGCAATTTGTTCCTCACTACACACCAACGACGACGATCATATATCGATTG	1191
DB	352	LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu	371
QY	1192	AGGGAGCGGCTCAGTGCAGTCTGGACAGCAGCAGGCAACAGAGTGTACGACGAGG	1251
DB	372	AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu	391
QY	1252	CTTCAAGAGGTCACGTCTGTGGTCCACAGAACTTCGCGTCCGCTGGAAGTCCAG	1311
DB	392	LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln	411
QY	1312	ACGAGAACTTCGAATACGTGGCATTCAGACAGACTCAAGGCCACGATAGCAACCTC	1371
DB	412	SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu	431
QY	1372	GCGGTGAAACTCCGTCATAGTAACCTGCGGAGGAGGTGTTGCAATTCATATGCG	1431
DB	432	AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn	451
QY	1432	CTCCAAGGAGCAGGCAAGCGAGCTTAAGAACACCAACCCCTTCAAGTCTTCGTTCCA	1491
DB	452	LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro	471
QY	1492	CCGTCTCAGCAGTCTCCGAGGCGTGTGCT 1521	
DB	472	ProLysGluSerGlnArgArgValValAla 481	

## RESULT 15

ABG71265  
 ID ABG71265 standard; protein; 485 AA.

XX  
 AC ABG71265;

XX  
 DT 17-DEC-2002 (first entry)

XX  
 DE Glycine max (Soybean) var. Shiroturunoko protein #2.

XX  
 KW Soybean; Glycinin; atomic coordinate data; processability; soya protein;  
 KW Shiroturunoko; protein co-ordinate data.

XX  
 OS Glycine max.

XX  
 PN JP2002193996-A.

XX  
 PD 10-JUL-2002.



PF 21-DEC-2000; 2000JP-00405097.  
XX PR 21-DEC-2000; 2000JP-00405097.  
XX PA (KYOU ) UNIV KYOTO.  
XX WPI; 2002-685438/74.  
XX DR N-PSDB; ABS55192.  
XX Glycinin, beta-conglycinin and proglycinin, their crystal structures,  
PT three dimensional coordinates, three dimensional structured and models  
PT and their uses.  
XX Disclosure; Page 1269-1271; 1298pp; Japanese.  
XX The present invention relates to a new Glycinin characterised by the  
CC atomic coordinate data fully defined in the specification. The structure  
CC can be used for improving processability of soya protein. The present  
CC amino acid sequence represents the Glycine max (Soybean) var.  
CC Shiotsurunoko protein #2, as described in the specification  
XX SQ Sequence 485 AA;  
  
Alignment Scores:  
Pred. No.: 2,25e-128 Length: 485  
Score: 1499.50 Matches: 297  
Percent Similarity: 72.02% Conservative: 71  
Best Local Similarity: 58.12% Mismatches: 94  
Query Match: 55.72% Indels: 49  
DB: 5 Gaps: 6  
  
US-10-728-051-3 (1-1524) x ABG71265 (1-485)  
QY 1 CGGCAGCAACGGAGGAGAGCGGTGCGAGTTCAGCGCCTCAATGCGAGACCTGAC 60  
DB 20 ArgGluGlnAlaGlnGlnGluCysGlnIleGlnLysLeuAenAlaLeuLysProAsp 39  
QY 61 AATCGCATTAATCAGAGGCGGTATCATTTAGACTTTGGAAACCCCAACACGAGGATTC 120  
DB 40 AsnArgileGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnLysProPhe 59  
QY 121 GAATGCGCGCGCGCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTCGTAGGCGCT 180  
DB 60 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 79  
QY 181 TTCTACTCCAAATGTCCTCCAGGAGATCTTCATCCAGCAAGAGGGGATCTTTGGGTTG 240  
DB 80 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 99  
QY 241 ATATTCCTGTTGTCCTAGACACTATGAAGAGCCTCACACAGGTCGTCATCTCAG 300  
DB 100 IlePheProGlyCysProSerThrTyrGlnGluProGlnGluSerGlnGlnArgGlyArg 119  
QY 301 TCCCAAGACCAACAGACGCTCCAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360  
DB 120 SerGlnArgPro-----GlnAspArg 126  
QY 361 CACCAAGGTGACCGCTTTCGATGAGGTGATCTCATTCAGTTCACCGCGGTGTTGCT 420  
DB 127 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 146  
QY 421 TTCTGGCTCTACAAACGACGACACTGATGTTGTTGTTGTTCTTTCTTACTGACCAAC 480  
DB 147 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 166  
QY 481 AACNACACACACCGCTTGATTCAGTTCCTCCAGGAGATTCATTTGGCTGGGACACGGAG 540  
DB 167 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186  
QY 541 CAAGAGTTCTTAAGGTACACGACCAACAGGAGAGCAAGAGCAAGAGGAGGATACCATAT 600  
DB 187 GlnGluPheLeuLysTyrGlnGlnGln----- 195

QY 601 AGCCCATACAGCCCGCAAGTCTAGCCTAGACAGAGCGCTGAATTTAGCCCTCGAGGA 660  
DB 196 -----GlnGlnGlyGlySerGlnSerGlnLysGly 205  
QY 661 CAGCACAGCGCAGAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
DB 206 LysGln-----GlnGluGluGluAsnGluGlySerAsnIleLeu 218  
QY 721 AGCGGTTTACGCGCGAGTTCCTGGAACAGCCCTTCAGGTTTACGACAGACAGATAGTG 780  
DB 219 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 237  
QY 781 CAAACCTTAAGGCGCAG 840  
DB 238 ArgAsnLeuGlnGlyGluAsnGluGluGluAsnSerGlyAlaIleValThrValLysGly 257  
QY 841 GGCTCAGAAATCTTGAAG 900  
DB 258 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluGluAspAspAsp 277  
QY 901 GAAGATGAA-----TATGAATAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948  
DB 278 GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 297  
QY 949 AGCAGAGGAGGCGGAGTATTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008  
DB 298 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 314  
QY 1009 ATTGGTAGAAACAGATCCCTGACATCTACAACCTCAAGCTGTTTCACTCAAAATGCC 1068  
DB 315 IleGlyGlnAsnSerSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 334  
QY 1069 AACGATCTCAACCTTCTAATATTAGTGGCTTGGACCTAGTCTGAATATGAATCTC 1128  
DB 335 ThrSerLeuAspPheProAlaLeuTyrLeuLysLeuSerAlaGlnTyrGlySerLeu 354  
QY 1129 TACAGAAATGCTTGTTCGCTCTACTACAAACCAACGACACAGCATCATATATCGA 1188  
DB 355 ArgLysAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 374  
QY 1189 TTGAGGGGACGGCTCAGCTGCAAGTCGTGACAGCAACGGCAACAGAGTGTACGACGAG 1248  
DB 375 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 394  
QY 1249 GAGCTTCAAGAGGTCAGCTGTTGTTGGTCCACAGAACTTCGCGTCTGCGAAAGTCC 1308  
DB 395 GluLeuGlnGlyGlyValLeuIleValProGlnAsnPheAlaValAlaLysSer 414  
QY 1309 CAGAGCGAGAACTTCGAATACGTGGCATTCAGACAGACTCAAGCCCGCAGCATAGCCAAC 1368  
DB 415 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434  
QY 1369 CTGCGCGTGAAACTCCGTCATAGATAACTCCGAGGAGGTGGTTGCAAAATTCATAT 1428  
DB 435 LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 454  
QY 1429 GGCTTCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1488  
DB 455 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 474  
QY 1489 CACCGCTCTCAGCAGTCTCGAGGGCTGGCT 1521  
DB 475 ProProGlnGluSerGlnArgArgAlaValAla 485

Search completed: April 8, 2005, 05:01:55  
Job time : 238.83 secs

**This Page Blank (uspto)**



Best Local Similarity: 100.00%		Mismatches: 0	
Query Match: 99.33%		Indels: 0	
DB: 2		Gaps: 0	
US-10-728-051-3 (1-1524) x 082580 (1-507)			
QY	1	CGGCAGCAACCGGAGGAGAACGGTCCAGCTTCCAGCGCTCAATGCGGAGAGACCTGAC	60
DB	1	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	20
QY	61	AATCGCAATGAATCAGAGGGGGTTCATGACATTTGGAACCCCAACACGAGAGTTC	120
DB	21	AsnArgileGluSerGluGlyGlyTyrIleGluThrTyrAsnProAsnAsnGlnGluPhe	40
QY	121	GAATGCGGGGCTGCGCTCTCTGCTTAGTCTCCGCGGCAACGCCCTTGGTAGGCT	180
DB	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	60
QY	181	TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATCTTTGGGTTG	240
DB	61	PheTyrSerAsnAlaProGlnGlnIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	80
QY	241	ATATTCCTGGTGTCTAGACACTATGAAGAGCTTCACACAAAGTGTGATCTCAG	300
DB	81	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln	100
QY	301	TCCCAAGACCAACAGAGCTCTCCAGGAGAGAACCAAGCCCAACAGCAACGAGATCT	360
DB	101	SerGlnArgProProArgLeuGlnGlnGlyGluAspGlnSerGlnGlnArgAspSer	120
QY	361	CACGAGAAGTGCACCGTTCGATGAGGTGATCTCATTCGATTCCTCCACCGGTGTGCT	420
DB	121	HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	140
QY	421	TTCTGGCTTACCAACGACGACACTGATGTTGTTGCTGTTCTTCTTACTGACCAAC	480
DB	141	PheTyrLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	160
QY	481	AACAAAGCAACCAAGCTGATGATGTTCCCGAGAGATTCATTTGCTGGGACACGGAG	540
DB	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	180
QY	541	CAAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCAT	600
DB	181	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	200
QY	601	AGCCCATACGCGCGCAAGTACCTAGCTCCCGAGAGATTCATTTGCTGGGACACGGAG	660
DB	201	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	220
QY	661	CAGCAGCGCGCAGAGAACGAGCAGACAGCAAGAGAGAGAAACCAAGGTGGAACATCTTC	720
DB	221	GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGlyAsnIlePhe	240
QY	721	AGCGGTTACGCGGAGTTCCTGGAACAAGCCCTTCAGGTTGACACAGACAGATAGTG	780
DB	241	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspArgGlnIleVal	260
QY	781	CAAAACCTAGAGCGGAGCGGAGTGTAGAGAGGAGGACCTTCTGACAGTGGGGA	840
DB	261	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	280
QY	841	GGCTCAGATCTTACGCCAGATAGAAAGAGACGTCGCCGACCAAGAGAGGAATACGAT	900
DB	281	GlyLeuArgileLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp	300
QY	901	GAGATGAATATGAATACGATGAGGATAGAGGCTGGCAGGGGAGGAGCAGGAGG	960
DB	301	GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg	320
QY	961	GGGATGGTATGAGAGAGCTGTCACCGCAAGTCTAAAAGACATTTGCTAGAAAC	1020
DB	321	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	340

Query Match:	92.12%	Indels:	4
DB:	2	Gaps:	2
US-10-728-051-3 (1-1524) x Q9F211 (1-529)			
QY	1	CGCAGCAACCGGAGGAGACGGTGCAGTTCAGCGCCTCAATGCGCAGACCTGAC	60
DB	25	ArgGlnGlnProGluGluAenAlaCysGlnPheGlnArgLeuAenAlaGlnArgProAsp	44
QY	61	AATCGCATTAATGATCAGAGGCGGTTCATTGAGACTTTGGAAACCCCAACACACGAGGATTC	120
DB	45	AsnArgLeuGluSerGluGlyGlyTyrIleGluThrTrpAsnProAenAenGlnGluPhe	64
QY	121	GAATGCGCGCGTGCCTCTCTGCTTAGTCTCTCGCGCGCAACGCGCTTCGTAGCCT	180
DB	65	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAenAlaLeuArgArgPro	84
QY	181	TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCCAGCAAGGAAGGGATATCTTGGGTG	240
DB	85	PheTyrSerAenAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	104
QY	241	ATATTCCCTGCTGCTCCTAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCAG	300
DB	105	IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgA-GHisGln	124
QY	301	TCCCAAGACCAACAGAGCTCTCCAAAGGAGAGACCAAGC---CAACAGCAACGAGAT	357
DB	125	SerGlnArgAlaProArgArgPheGluGlyGluAspGlnSerGlnGlnGlnGlnAsp	144
QY	358	ASTCACCAGAGTGCACCTTTCGATGAGGTGATCTCATTGCAGTCTCCACCGGTGT	417
DB	145	SerHisGlnIysValArgArgPheAspGluGlyAspLeuIleAlaValProThrGlyVal	164
QY	418	GCTTTCTGGCTCTACACAGCACGACACTGATGTTGCTGCTGTTCTTCTTACTGACAC	477
DB	165	AlaLeuTrpMetPheAenAspHisAspThrAspValValAlaValSerLeuThrAspThr	184
QY	478	AACAACCAACCAACAGCTTGATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACAG	537
DB	185	AsnAenAenAenGlnLeuAenGlnPheProArgArgPheAenLeuAlaGlyAenHis	204
QY	538	GACCAAGATCTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGACGACGACCTTACCA	597
DB	205	GluGlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuPro	221
QY	598	TATAGCCCATACAGCCGCAAGTACGCTAGACAAAGAGAGTTCATTTGGCTGGGAACAG	657
DB	222	TyrSerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArg	241
QY	658	GGACAGCACCGCCAGCAAGACGAGCAAGAGCAAGAAAGAAACGAAGTGGAAACATC	717
DB	242	GlyGlnHisSerArgArgGluArgAlaGlyGlnGluGluAenGluGlyGlyAenIle	261
QY	718	TTCAGCGCTTCAGCCGGAGTTCCTGGAACAGCTTCAGGTTCAGCAGCAGACAGATA	777
DB	262	PheSerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIle	281
QY	778	GTGCAAAACCTTAAGCGCAGACCGCAGAGTGAAGAGGAGGAGCTTGTGACAGTGAAG	837
DB	282	ValGlnAenLeuArgGlyGluAenGluSerGluGluGlyAlaIleValThrValIys	301
QY	838	GGAGGCTTCAGAACTTTGAGCCAGATAGAAAGAGAGCTGCCGACGAAGAGGAATAC	897
DB	302	GlyGlyLeuArgIleLeuSerProAspArgIysArgGlyAlaAspGluGluGluGluTyr	321
QY	898	GATCAAGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	957
DB	322	AspGluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGly	341
QY	958	AGGGGGAATGATTTGAGAGACCATCTGCACCGCAAGTGTCTAAAGAAACATTTGGTGA	1017
DB	342	ArgGlyAenGlyIleGluGluThrIleCysThrAlaSerValIysIysAenIleGlyArg	361
QY	1018	AACAGATCCCTGACATCTACACCTCAAGCTGGTTCCTCACTCAAACTGCCACGATCTC	1077
DB	362	AsnArgSerProAspIleTyrAenProGlnAlaGlySerLeuIysThrAlaAenAspLeu	381
QY	1078	AACCTTCTAATACTAGTGGCTTGACCTAGTCTGAATATATGAAATCTCTACAGAAAT	1137
DB	382	AsnLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAenLeuTyrArgAen	401
QY	1138	GAATTTTGTTCGCTCATTACCAACCAACGACACAGCATCATATATCGATTGAGGGA	1197
DB	402	AlaLeuPheValProHisTyrAenThrAenAlaHisSerIleIleTyrAlaLeuArgGly	421
QY	1198	CGGGCTCAGTCCCAAGTCTGACGACCAACGCAACAGAGTGTACACGAGGAGCTTCAA	1257
DB	422	ArgAlaHisValGlnValValAspSerAenGlyAenArgValTyrAspGluGluLeuGln	441
QY	1258	GAGGGTCAGTGTCTGCTGGTGCACAGAACTTCGCGCTCGCTGAAAGTCCACGAGCGAG	1317
DB	442	GluGlyHisValLeuValValProGlnAenPheAlaValAlaGlyIysSerGlnSerAsp	461
QY	1318	AATCTCGAATACGTGTCATTCAGACAGACTCAAGCCCGCAGCATGCCAACCTTCGCGGT	1377
DB	462	AsnPheGluTyrValAlaPheIysThrAspSerArgProAenIleAlaAenPheAlaGly	481
QY	1378	GAATACTCGTCATAGATAACCTCCGCGAGGAGGTGTTGCCAATTCATATGCGCTCAA	1437
DB	482	GluAenSerIleIleAspAenLeuProGluGluValValAlaAenSerTyrGlyLeuPro	501
QY	1438	AGGGAGCAGCAGGAGGAGCTTAAAGAACAAACACCCCTCAAGTCTTCTGTTCCACCGTCT	1497
DB	502	ArgGluGlnAlaArgGlnLeuIysAenAenAenProPheIysPheValProProSer	521
QY	1498	CAGCAGTCTCCGAGGCTGTGGCT 1521	
DB	522	GlnGlnSerLeuArgAlaValAla 529	
RESULT 3			
ID	Q647H4	PRELIMINARY;	PRT; 536 AA.
AC	Q647H4;		
DT	25-OCT-2004	(TrEMBLrel. 28, Created)	
DT	25-OCT-2004	(TrEMBLrel. 28, Last sequence update)	
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)	
DE	Arachin Ahy-1.		
OS	Arachis hypogaea (peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Yan Y., Wang L., Huang S.;		
RT	"cDNA clone of peanut seed storage protein gene";		
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; A772685; AAU21490.1; -.		
SQ	SEQUENCE 536 AA; 61505 MW; 30DBF638719AEB78 CRC64;		
Alignment Scores:			
Pred. No.:	1,02e-163	Length:	536
Score:	2455.00	Matches:	471
Percent Similarity:	93.99%	Conservative:	14
Best Local Similarity:	91.28%	Mismatches:	19
Query Match:	91.23%	Indels:	12
DB:	2	Gaps:	3
US-10-728-051-3 (1-1524) x Q647H4 (1-536)			
QY	1	CGCAGCAACCGGAGGAGACGGTGCAGTTCAGCGCCTCAATGCGCAGACCTGAC	60
DB	24	ArgGlnGlnProGluGluAenAlaCysGlnPheGlnArgLeuAenAlaGlnArgProAsp	43
QY	61	AATCGCATTAATGATCAGAGGCGGTTCATTGAGACTTTGGAAACCCCAACACGAGGATTC	120

```
Db 44 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
QY 121 GAATGGCGGGCGCTCTCTCGCTTAGTCTCCGCGGAAAGCCCTTCGTAGGCGCT 180
Db 64 GluCyAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 83
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGTTG 240
Db 84 PheTyrSerAsnAlaProGlnGluPheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
QY 241 ATATTCCTCGTTGCTCTAGACACTATGAAGAGCCTCACACAGGTCGTGATCTCAG 300
Db 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgGlyHisGln 123
QY 301 TCCCAAGACCAACAGAGCTCTCCAGGAGAGAGCAAGCCCAACAGCAACAGATAGT 360
Db 124 SerGluArgProArgArgPheGlnGlyGlnAspGlnSerGlnGlnGlnGlnAspSer 143
QY 361 CACCAGAAGTGCA CCGTTTCGATGAGGAGTCTCTCATTCGAGTTCACACCGGTGTGCT 420
Db 144 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 163
QY 421 TTCGTGCTCTAACACACAGCAGCTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
Db 164 PheTrpMetTyrAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn 183
QY 481 AACAAACGACACCGCTTGATGATGTTCCCGAGGAGATTCATTTGCTGGGAACACGGAG 540
Db 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 203
QY 541 CAAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGAGCAAGAGCTTACCATAT 600
Db 204 GlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuProTyr 220
QY 601 AGCCCATACAGCCGCAAGTACGCTAGCAGAGAGAGCGTGAATTTAGCCCTCCAGAG 660
Db 221 SerProTyrSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly 240
QY 661 CAGCAGACCGCAGAGAACGAGCAGGAGCAAGAGAGAGAAACCAAGCTGGAACATCTTC 720
Db 241 GlnHisGlyArgArgGluArgAlaGlyGlnGlnGluGlnGluAsnGluGlyGlyAsnIlePhe 260
QY 721 AGCGGCTTACGCGGAGTCTCTGGAAACAAGCTTCAGGTTCCAGCAGCAGACAGATAGTG 780
Db 261 SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIleLeu 280
QY 781 CAAACCTAAGAGCGGAGCCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGGA 840
Db 281 GlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArgGly 300
QY 841 GGCCTCAGAAATCTTGAGCCAGATAGAAAGAGA-----CGTGCAGCAGAA 885
Db 301 GlyLeuArgIleLeuSerProAspArgLysArgGlnGlnTyrGluArgProAspGlu 320
QY 886 GAAGAGGAATACGATGAAGATGAATATGAATACATGATGAAG-----GATAGA 933
Db 321 GluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluArgGlnHisAspArg 340
QY 934 AGCGGTGGCAGGGGACGAGCAGCGGGAATGTTATTCAGACAGCAGATCTGACCGCA 993
Db 341 ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla 360
QY 994 AGTGCTAAAAAGAACATTTGTTAGAAAACAGATCCCTTCACATCTACAAACCTCAAGCTGGT 1053
Db 361 SerPheLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAlaGly 380
QY 1054 TCACTCAAACTGCAACGATCTCAACCTTCTAATCTAGGTGGCTTGACACCTAGTGCT 1113
Db 381 SerLeuLysThrAlaAsnGluLeuAsnLeuIleLeuArgTrpLeuGlyLeuSerAla 400
QY 1114 GAATATGGAATCTCTACAGGAATGCAATGTTGTGCTCACTACACACACACGACAC 1173
```

```
Db 401 GluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAlaHis 420
QY 1174 AGCATCATATATCGATTCAGGGGACGGGCTCACGTGCAAGTCGTGAGCAGCAACGGCAAC 1233
Db 421 SerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGlyAsp 440
QY 1234 AGAGTGACAGCAGGAGGCTTCAAGAGGTCACGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1293
Db 441 ArgValPheAspGluGluGlnGluGlyHisValLeuValValProGlnAsnPheAla 460
QY 1294 GTCGCTGGAAGTCCAGAGCGAGAACTTCCAATACGTGGCATTCAGACAGACACTCAAGG 1353
Db 461 ValAlaGlyLysSerGlnSerGluAsnPheGluTyrValAlaPheLysThrAspSerArg 480
QY 1354 CCCAGCATAGCAACCTCCCGGTGAAAACTCCGTCATAGATAAACCTCCCGGAGAGGTG 1413
Db 481 ProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAspAsnLeuProGluGluVal 500
QY 1414 GTTGAATTCATATGCGCTCCAAAGGAGCAGGAGCAGGCTTAAAGCAACACCAACCCC 1473
Db 501 ValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsnAsnPro 520
QY 1474 TTCAAAGTTCCTCGTTCCACCGCTCTCAGCAGTCTCCGAGGCTGTGGCT 1521
Db 521 PheLysPhePheValProSerGluGlnSerLeuArgAlaValAla 536
RESULT 4
Q647H3 PRELIMINARY; PRT; 537 AA.
AC Q647H3;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE Arachin Ahy-2.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722686; AAU21491.1; -.
SQ SEQUENCE 537 AA; 61531 MW; 3BDD034DEA159657 CRC64;
Alignment Scores:
Pred. No.: 1,8e-163 Length: 537
Score: 2451.50 Matches: 474
Percent Similarity: 93.62% Conservative: 10
Best Local Similarity: 91.68% Mismatches: 20
Query Match: 91.10% Indels: 13
DB: 2 Gaps: 4
US-10-728-051-3 (1-1524) x Q647H3 (1-537)
```

```
QY 1 CGGCAGCAACCGGAGGAGAACCGTCCAGTTCAGCGCTCAATCGCAGAGACTGAC 60
Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 ATTCGCAATGAATCAGAGCGCGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
Db 44 AsnArgLeuGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
QY 121 GAATGGCGGGCGTCCGCTCTCTCGTTAGTCTCCGCGCAACGCCCTTCGTAGGCT 180
Db 64 GluCyAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 83
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGTTG 240
Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
```

QY 241 ATATTCCCTGGTGTCTCTAGACACTATGAAGAGCCCTCACACAACAGGTGCTCGATCTCAG 300  
Db |||||  
QY 104 IlePheProGlyCysProSerThrTyrgluGluProAlaGlnGlnGlyArgArgHisGln 123  
Db |||||  
QY 301 TCCCAAGACCCCAAGAGCTCTCCAAAGGAGAGACCAAGC--CAACAGACACAGAT 357  
Db |||||  
QY 124 SerGlnArgAlaProArgArgPheGluGlyGluSerGlnGlnGlnGlnGlnGln 143  
Db |||||  
QY 358 AGTCACCAAGGTGCACCGCTTTCGATGAGGGTGATCTCATTCAGTCTCCACCGGTGTT 417  
Db |||||  
QY 144 SerHisGlnIysValArgArgPheAspGluGlyAspLeuIleAlaValProThrGlyVal 163  
Db |||||  
QY 418 GCTTTCTGGCTCTACACGACCAAGACACGACATGATGTTGTGCTGCTTCTTACTGACACC 477  
Db |||||  
QY 164 AlaLeuTrpMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThr 183  
Db |||||  
QY 478 AACCAACAGACACACAGCTTGATCTGATCTCCAGGAGATTCATTTGGCTGGACAGC 537  
Db |||||  
QY 184 AsnAsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHis 203  
Db |||||  
QY 538 GAGCAAGAGTCTTAAAGGTACACAGCAACAAAGCAGACAAAGCAGACGAAGCTTACCA 597  
Db |||||  
QY 204 GluGlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuPro 220  
Db |||||  
QY 598 TATAGCCCATACAGCCCGCAAGTCTAGCTAGCAAGAGAGCGTGAATTTAGCCCTCGA 657  
Db |||||  
QY 221 TyrSerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArg 240  
Db |||||  
QY 658 GGCACACAGCCGACAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 717  
Db |||||  
QY 241 GlyGlnHisSerArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGln 260  
Db |||||  
QY 718 TTCAGCGGCTTTCAGCGGAGTCTCTGGAACAGCCTTCAGGTTGACACACACAGATA 777  
Db |||||  
QY 261 PheSerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIle 280  
Db |||||  
QY 778 GTGCAAAACCTAAGCGCAGACCGAGAGTGAAGAAGAGAGGAGCCATTTGACAGTGAGG 837  
Db |||||  
QY 281 LeuGlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlnGlyAlaIleValThrValArg 300  
Db |||||  
QY 838 GAGGCTCAGATCTTGACCCAGACAGACAGACAGACAGACAGACAGACAGACAGAC 882  
Db |||||  
QY 301 GlyGlyLeuArgIleLeuSerProAspArgLysArgArgGlnGlnTyrGluArgProAsp 320  
Db |||||  
QY 883 GAAGAAGAGGAATACGATGAAGATGAATATGATGATGATGATGATGATGATGATGAT 930  
Db |||||  
QY 321 GluGluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluGluGlnGlnAsp 340  
Db |||||  
QY 931 AGAAGCGTGGCAGGGGAAGCAGAGCGAGGGGAATGGTATTCAAGAGACGATCTGCACC 990  
Db |||||  
QY 341 ArgArgGlyArgGlySerArgGlyArgGlyAsnGlyIleGluGluThrIleCysThr 360  
Db |||||  
QY 991 GCAAGTGTCTAAAGACATGTTGTAGAACAGATCCCTGACATCTACACCCCTCAAGCT 1050  
Db |||||  
QY 361 AlaSerValIysIysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAla 380  
Db |||||  
QY 1051 GGTTCCTCAAACTGCCAAGATCTCAACCTTCTAATCTAGTGGCTTGGACCTAGT 1110  
Db |||||  
QY 381 GlySerLeuIysThrAlaAsnAspLeuAsnLeuIleLeuArgTrpLeuGlyLeuSer 400  
Db |||||  
QY 1111 GCTGAATATGAAATCTCTACAGGAATGCAATGTTGTCCTCTCACTACCAACCAACGCA 1170  
Db |||||  
QY 401 AlaGluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAla 420  
Db |||||  
QY 1171 CACAGATCATATATCGATTGAGGAGCGGGCTCAGCTGCAAGTCTGTGACAGACAGGC 1230  
Db |||||  
QY 421 HisSerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGly 440  
Db |||||  
QY 1231 AACAGAGTGTACCGAGGAGCTTCAAGAGGTCAGTGCTGCTGTCGTCACAGACTTC 1290  
Db |||||  
QY 441 AsnArgValTyrAspGluGluLeuGlnGlnGlyHisValLeuValValProGlnAsnPhe 460

QY 1291 GCGTCGCTGGAAAGTCCAGAGCGAGAACTTCGAATATCGTGCATTCAAGACAGACTCA 1350  
Db |||||  
QY 461 AlaValAlaGlyLysSerGlnSerAspAsnPheGluTyrValAlaPheLysThrAspSer 480  
Db |||||  
QY 1351 AGCCCCAGCATAGCCAACTCCCGGTGAAACTCCGTGCATAGATAACCTGCCGAGGAG 1410  
Db |||||  
QY 481 ArgProSerIleAlaAsnLeuAlaGlyGluAsnSerIleIleAsnLeuProGluGlu 500  
Db |||||  
QY 1411 GTGGTTGCAAAATTCATATGCTCTCCAAAGGAGCAGCAGCAGCAGCAGCAGCAGC 1470  
Db |||||  
QY 501 ValValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 520  
Db |||||  
QY 1471 CCCTTCAAGTCTTCTGTTCCACCGTCTCAGCAGTCTCCGAGGCTGTGGCT 1521  
Db |||||  
QY 521 ProPheLysPhePheValProSerGlnGlnSerLeuGlyAlaValAla 537  
Db |||||  
RESULT 5  
O8LKN1 PRELIMINARY; PRT; 538 AA.  
ID Q8LKN1  
AC Q8LKN1  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Allergen Arab3/Arab4.  
OS Arachis hypogaea (peanut).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
OC Arachis.  
OX NCBI\_TaxID=3818;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Viquez O.M., Konan K.N., Dodo H.W.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF510854; AAM46958.1; -.  
DR HSSP; P04776; IUCX.  
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
DR InterPro; IPR006045; Cupin.  
DR InterPro; IPR007113; Cupin region.  
DR InterPro; IPR011051; RmlC like cupin.  
DR InterPro; IPR006044; Seedstore\_115.  
DR Pfam; PF00190; Cupin; 2.  
DR PRINTS; PR00439; 11SGLOBULIN.  
DR PROSITE; PS00305; 11S SEED STORAGE; UNKNOWN 1.  
SQ SEQUENCE 538 AA; 61737 MW; 7AABD0D59429709E CRC64;  
Alignment Scores:  
Pred. No.: 6,04e-163 Length: 538  
Score: 2444.00 Matches: 471  
Percent Similarity: 93.63% Conservative: 14  
Best Local Similarity: 90.93% Mismatches: 19  
Query Match: 90.82% Indels: 14  
DB: 2 Gaps: 4  
US-10-728-051-3 (1-1524) x Q8LKN1 (1-538)  
QY 1 CGCGACCAACCGGAGAGAACCGTCCAGTTCAGCGCTCAATCGCAGACCTGAC 60  
Db |||||  
QY 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43  
Db |||||  
QY 61 AATCGCATTAATACAGAGCGGCTTACATGAGACTTGGAAACCCCAACCAAGGAGTTC 120  
Db |||||  
QY 44 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnGlnGluPhe 63  
Db |||||  
QY 121 GAATGCGCGGCTGCGCTCTCTCGCTTAGTCCTCGCGGCAACCGCTTCGTAGCCT 180  
Db |||||  
QY 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgPro 83  
Db |||||  
QY 181 TTCTACTCAATAGTCTCCCGAGGAGATCTTTCATCCAGCAAGGAGGATATCTTTGGTTG 240  
Db |||||  
QY 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103  
Db |||||  
QY 241 ATATTCCCTGGTGTCTCTAGACACTATGAAGAGCCTCACACAACAGGTGCTCGATCTCAG 300



```
Db 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgHisGln 123
|||
QY 301 TCCAAAGACCCAAAGAGCTCTCCAAAGGAGAGACCAAAGGACCAACAGACGAGTACT 360
|||
Db 124 SerGlnArgProProArgArgPheGlnGlyGlnAspGlnSerGlnGlnGlnAspSer 143
|||
QY 361 CACCAGAGGTGACCGCTTCATGATGAGGTGATCTCATTTGCAGTTCCTCCACCGGTGTGCT 420
|||
Db 144 HisGlnLysValHisArgPheAspGluGlyAspLeuLeuAlaValProThrGlyValAla 163
|||
QY 421 TTCTGGCTCTACAAACACACACACACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
|||
Db 164 PheTrpMetCysAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
|||
QY 481 AACACGACCAACCCAGCTTGATCAGTTCCCCAGGAGATTCAATTTGCTGGGAACACGAG 540
|||
Db 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 203
|||
QY 541 CAAGATCTTAAAGTACCAGCAACAAAGCAGACGAAAGCAGACGAAAGAGCTTACCATAT 600
|||
Db 204 GlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuProTyr 220
|||
QY 601 AGCCCATACAGCCCGAAGCTCAGCTAGACAGAGAGCGTCAATTTAGCCCTCGAGGA 660
|||
Db 221 SerProTyrSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly 240
|||
QY 661 CAGCAGACCGCAGACAGACGAGCAGACGAGCAAGAGAAAGAAAGCAAGCTGGAACATCTTC 720
|||
Db 241 GlnHisGlyArgArgGluArgAlaGlyGlnGlnGluAsnGluGlyGlyAsnIlePhe 260
|||
QY 721 AGCGGTCTACCGCGAGTCTCTGGAAACAAGCTTCCAGGTTGACACAGACAGATAGTG 780
|||
Db 261 SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIleLeu 280
|||
QY 781 CAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGCCATTGTGACGTAGAGGGA 840
|||
Db 281 GlnAsnLeuArgGlyGlnAsnGluSerAspGluGlnGlyAlaValThrValArgGly 300
|||
QY 841 GGCTCTCAGAACTTTGAGCCCGACAGTAAAGAGAG-----CGTCCCGACGAA 885
|||
Db 301 GlyLeuArgIleLeuSerProAspArgLysArgArgGlnGlnTyrGluArgProAspGlu 320
|||
QY 886 GAAGAGGATACATGAAGATGAATATGATAGCATGAAG-----GATAGA 933
|||
Db 321 GluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluGluArgGlnGlnAspArg 340
|||
QY 934 AGCGTGGCAGGGGAGCAGAGCGAGGGAATGGTATTGAAGAGACGATCTGCACCGCA 993
|||
Db 341 ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla 360
|||
QY 994 AGTGCTAAAGAACATTTGGTAGAAACAGATCCCTGACATCTACAACTCAAGCTGGT 1053
|||
Db 361 SerPheLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAlaGly 380
|||
QY 1054 TCACTCAAACTCCCAACGAT-----CTCAACTCTTAATCTTAGTGGCTTGACCT 1107
|||
Db 381 SerLeuLysThrAlaAsnGluLeuGlnLeuAsnLeuLeuLeuArgTrpLeuGlyLeu 400
|||
QY 1108 AGTGCTGAATATGGAATCTTACAGGAATGCATTGTTGTGCTCACTACACACCAAC 1167
|||
Db 401 SerAlaGluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsn 420
|||
QY 1168 GCACACAGCATCATATATGATTGAGGGGACGGGCTCAGGTGCAAGTCTGGACAGCAAC 1227
|||
Db 421 AlaHisSerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsn 440
|||
QY 1228 GGCACAGAGTGTACAGCAGGAGCTTCAAGAGGTCAGCTGCTGTGGTCCACAGAAC 1287
|||
Db 441 GlyAspArgValPheAspGluGluLeuGlnGluGlyHisValLeuValValProGlnAsn 460
|||
QY 1288 TTCGCGTGTGCTGGAAGTCCACAGAGCGAAGCTTCGAATACGTGCATTCAAGACAGAC 1347
|||
```

```
Db 461 PheAlaValAlaGlyLysSerGlnSerGluAsnPheGluTyrValAlaPheLysThrAsp 480
|||
QY 1348 TCAAGCCGACGATAGCCAACTCCCGGTGAAACTCCGTCTATAGATAAACCCTCCGGAG 1407
|||
Db 481 SerArgProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAspAsnLeuProGlu 500
|||
QY 1408 GAGGTGGTTGCAAAATTCATATGGCTCCAAAGGAGCAGGCAAGCAGCTTAAAGAACAC 1467
|||
Db 501 GluValValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 520
|||
QY 1468 AACCCCTTCAAGTCTTCTGTTCCACCGTCTCAGCAGTCTCCAGGGCTGTGGCT 1521
|||
Db 521 AsnProPheLysPhePheValProProSerGluGlnSerLeuArgAlaValAla 538
|||
RESULT 6
Q6T2T4 PRELIMINARY; PRT; 536 AA.
AC Q6T2T4;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Storage protein.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang H.-X., Wang F., Bi Y.-P.;
RL Submitted (Oct-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY439332; AAR02860.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmC_Like_cupin.
DR InterPro; IPR006044; Seedstore_lls.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S SEED STORAGE; UNKOWN 1.
SQ SEQUENCE 536 AA; 61498 MW; 138F3C881BC2C7D5 CRC64;

Alignment Scores:
Pred. No.: 2,2e-162 Length: 536
Score: 2436.00 Matches: 468
Percent Similarity: 93.60% Conservative: 15
Best Local Similarity: 90.70% Mismatches: 21
Query Match: 90.52% Indels: 12
DB: 2 Gaps: 3

US-10-728-051-3 (1-1524) x Q6T2T4 (1-536)
QY 1 CGGCAGCAACCGGAGGAGAACCGTGCACAGTTCACAGCGCTCAATCGCGCAGAGACCTGAC 60
|||
Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
|||
QY 61 AATCGCATTAATCAGAGGCGGTTTACATTGAGACTTGGAAACCCCAACAAACAGAGGTTTC 120
|||
Db 44 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnGlnGluPhe 63
|||
QY 121 GAATGGCGCGGCTGCCCTCTCTCGTTAGTTCCTCCCGCGCAACGCCCTTCTGTAGGCT 180
|||
Db 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 83
|||
QY 181 TTCTACTCTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGGTGG 240
|||
Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgAlaTyrPheGlyLeu 103
|||
QY 241 ATATTCCCTGGTGTCTTAGACACTATGAAGAGCTTCACACAAAGTCTCGATCTCAG 300
|||
Db 104 IlePheLeuGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgHisGln 123
|||
```

```
QY 301 TCCCAAGACACCAAGACCTCTCCAGGAGAGAGACCAAGACCAAGACGAGTAGT 360
Db 124 SerGlnArgProProArgPheGlnGlyAspGlnSerGlnGlnGlnGlnAspSer 143
QY 361 CACCAAGGTGACACCGCTTCGATGAGGGTGATCTCATTCAGCTTCCACCGGTGTGCT 420
Db 144 HisGlnLysValHisArgPheAspGluGlyAspLeuLeuValProThrGlyValAla 163
QY 421 TTCTGGCTCTACACGACGACGACATGATGTGTGCTGCTTCTTCTTACTGACACCAAC 480
Db 164 PheTrpMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
QY 481 AACCAACACACACGCTTGTATGATTCCTCCAGGAGATTCAATTTGGCTGGGAAACCGGAG 540
Db 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHisGlu 203
QY 541 CAAGAGTCTTAAAGTACACGACCAACAAAGCAGACAAAGCAGACGAGAGAGCTTACCATAT 600
Db 204 GlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuProTyr 220
QY 601 AGCCCATACACCGCGGAAAGTACGCTAGCCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db 221 SerProTyrSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly 240
QY 661 CAGCAGACCGCGCAGACGACGAGCAGACGAGACGAGACGAGACGAGTGGAAACATCTTC 720
Db 241 GlnHisGlyArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnPhe 260
QY 721 AGCGGCTTACCGCGGAGTCTCTGGAACAAAGCTTCCAGGCTTCCAGCAGACAGATAGTG 780
Db 261 SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnLeu 280
QY 781 CAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 281 GlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArgGly 300
QY 841 GGCCTCAGATCTTGACCCAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 885
Db 301 GlyLeuArgIleLeuSerProAspArgLysArgGlnGlnTyrGluArgProAspGlu 320
QY 886 GAAGAGGAATACGATGAAGATGAATATGAATACGATGAAGAGAGAGAGAGAGAGAGAG 933
Db 321 GluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluGluGlnGlnAspArg 340
QY 934 AGCGTGGCAGGGGAGACGAGAGCGGGGAGATGGTATTGAAGAGAGAGATCTCAGCGCA 993
Db 341 ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla 360
QY 994 AGTGCTAAAGAACATTTGCTAGAACAGATCCCTGACATCTACACCTCAGCTGCT 1053
Db 361 SerPheLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAlaGly 380
QY 1054 TCACTCAAACTGCCAACGATCAACCTTCTAATCTTAGGTGGCTTGGACCTAGTGCT 1113
Db 381 SerLeuLysThrAlaAsnGluLeuAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 400
QY 1114 GATATGGAATCTCTACAGGAATGCAATTTGCTGCTCCTACACACCAACAGCACAC 1173
Db 401 GluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAlaHis 420
QY 1174 AGCATCATATATCGATTGAGGAGCGGCTCAGTGCAGCTCGTGGACAGACAGCGCAAC 1233
Db 421 SerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGlyAsp 440
QY 1234 AGAGTGACACGAGGAGCTTCAAGAGGGTCACTGCTGCTGGTGGCCACAGAACTTCGCC 1293
Db 441 ArgValPheAspGluGluLeuGlnGlnGlyHisValLeuValValProGlnAsnPheAla 460
QY 1294 GTGCTGGAAGTCCAGAGCGGAGAACTTCGAATACGTGGCATTCAGACAGACTCAAGG 1353
Db 461 ValAlaGlyLysSerGlnSerGluAsnPheGluTyrValAlaPheLysThrAspSerArg 480
QY 1354 CCAGCATACCAACCTCGCGGGTGAATACTCCGTCTATAGTAACCTCGCGGAGGAGTG 1413
```

```
Db 481 ProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAsnLeuProGluGluVal 500
QY 1414 GTTGCATAATTCATATGCGCTCCAAAGGAGGAGCAGGAGCTTAAAGAACAAACCCC 1473
Db 501 ValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsnPro 520
QY 1474 TTCAGTCTTCTGTTCCACCGTCTCAGCAGTCTCCGAGGCGCTGGCT 1521
Db 521 PheLysPhePheValProProSerGluGlnSerLeuArgAlaValAla 536

RESULT 7
Q9SQH7 PRELIMINARY; PRT: 530 AA.
ID Q9SQH7 AC Q9SQH7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Glycinin.
GN Name=Arachis;
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Virginia;
RX MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;
RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.M.;
RT "Selective cloning of peanut allergens, including profilin and 2S
RL albumins, by phage display technology.";
RL Int. Arch. Allergy Immunol. 119:265-274(1999).
DR EMBL; AF086821; AAD47382.1; --
DR HSPB; P04776; IUCX.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR011051; RMLC_like_cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
SQ SEQUENCE 530 AA; 61011 MW; 864E349A31642A32 CRC64;

Alignment Scores:
Pred. No.: 2,48e-161 Length: 530
Score: 2421.00 Matches: 469
Percent Similarity: 94.09% Conservative: 9
Best Local Similarity: 92.32% Mismatches: 29
Query Match: 89.97% Indels: 2
DB: 2 Gaps: 0

US-10-728-051-3 (1-1524) x Q9SQH7 (1-530)
QY 1 CGGCACACCGGAGAGAGACCGCTGCCAGTTCACGGCTCAATGCGCGAGACCTGAC 60
Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCGCATTTGAATCAGAGGCGGTTCATTGACACTTGGAAACCCCAACACAGGAGTTC 120
Db 44 AsnArgIleGluSerGluGlyGlyIleGluThrIleGluThrIleGluThrIleGluThr 63
QY 121 GAATGCGCGGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGGCCT 180
Db 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 83
QY 181 TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGAGAGGGGATCTTTGGGTTG 240
Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
QY 241 ATATTCCCTGGTGTCTAGACACTATGAGAGCTTCACACACAGGCTCGTCGATCTCAG 300
```

```
Db 104 IlePheProGlyCysProSerThrTyrGluProAlaGlnGlnGlyArgArgTyrGln 123
QY 301 TCCCAAGACACCAAGACGTCTCCAGAGGAAGACCAAGCCAAAGCAGCAACGAGATAGT 360
Db 124 SerGlnArgProProArgArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 143
QY 361 CACCAGAGGTGACCGTTCGATGAGGTGATCTCATTCAGTCTCCACCGGTGTGCT 420
Db 144 HisGlnLysValHisArgPheAsnGluGlyAspLeuIleAlaValProThrGlyValAla 163
QY 421 TTCTGGCTCTACACGACGACGACGACTGATGTTGCTGCTTCTTCTTACTGACACCAAC 480
Db 164 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
QY 481 AACAAAGACACACGCTTGATCTCCAGGAGATTCATTTGGCTGGGACACGGAG 540
Db 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHisGlu 203
QY 541 CAAGAGTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT 600
Db 204 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 223
QY 601 AGCCCATACAGCCGCGAAAGTACGCTAGCAAGAGAGCGGTGAATTTAGCCCTCGAGGA 660
Db 224 SerProTyrSerProHisSerArgProArgArgGluGluGluPheArgProArgGly 243
QY 661 CAGCACAGCCGACAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 720
Db 244 GlnHisSerArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 263
QY 721 AGCGGTCTACGCGGAGTTCCTGGCAACAGCCCTCCAGGTTGACGACGACGACGATAGTG 780
Db 264 SerGlyPheThrProGluPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal 283
QY 781 CAAACCTAAGAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 840
Db 284 GlnAsnLeuTrpGlyGluAsnGluSerGluGluGluGluGluGluGluGluGluGluGly 303
QY 841 GGCTCAGAACTCTGAGCCAGATAGAACAGACGTCGCCAGCAAGAGAGGAATACGAT 900
Db 304 GlyLeuArgIleLeuSerProAspGlyThrArgGlyAlaAspGluGluGluTyrAsp 323
QY 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAAGCGTGGCAGGGAAGCAGACGAGG 960
Db 324 GluAspGlnTyrGlnTyrHisGluGlnAspGlyArgArgGlyArgGlySerArgGlyGly 343
QY 961 GCGAATGGTATTGAAGAGCAGTCTCCACCGCAAGTCTAAAGAACATTTGTAGAAC 1020
Db 344 GlyAsnGlyIleGluGluThrIleCysThrAlaCysValLysLysAsnIleGlyGlyAsn 363
QY 1021 AGATCCCTGACATCTACACCTCTCAAGC-TGGTTCACTCAAACTGCCAAACGATCTCAA 1079
Db 364 ArgSerProHisIleTyrAspProGlnArgTrpPheThrGlnAsnCys-HisAspLeuAs 383
QY 1080 CTTCTAACTACTAGTGGCTTGGACCTAGTGTGTAATATGGAATCTCTACAGGAATCG 1139
Db 383 nLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsnAl 403
QY 1140 ATTGTTTGCTCTACTACCAACCAACGACGACGACATCATATATCGATTGAGGGGAG 1199
Db 403 aLeuPheValProHisTyrAsnThrAsnAlaHisSerIleIleTyrAlaLeuArgGlyAr 423
QY 1200 GGCTCAGTCAGTGGTGGACAGCAACCGCAACAGAGGTGACGACGAGGAGCTTCAAGA 1259
Db 423 gAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnI 443
QY 1260 GGTCACGTCCTTGTGGTCCACAGAACTTCGCGCTCGCTGGAAAGTCCACAGCGAGAA 1319
Db 443 uGlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAs 463
QY 1320 CTTTCGAATACGTGCTTCAAGACAGACTCAAGGCCGACGATAGCAACTCGCGGTGA 1379
Db 463 nPheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnPheAlaGlyG 483
```

```
QY 1380 AAATCCGTCATAGATAACCTCCGAGGAGGTGTTGCAAAATTCATATGGCTCCAAAG 1439
Db 483 uAsnSerPheIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuProAr 503
QY 1440 GGAGCAGGCAAGCGAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCA 1499
Db 503 gGluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProPheG 523
QY 1500 GCAGTCTCCGAGGGCTGTGGCT 1521
Db 523 nGlnSerProArgAlaValAla 530

RESULT 8
Q61WG5 PRELIMINARY; PRT; 510 AA.
ID Q61WG5;
AC Q61WG5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Glycinin (Fragment)
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP SEQUENCE FROM N.A.
RA Kang I.-H., Gallo-Meagher M.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY618460; AAT39430.1; -.
DR GO; GO:0045735; P.nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR011051; RmlC_like_cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 510 AA; 58061 MW; F6569EF13D1A12BD CRC64;

Alignment Scores:
Pred. No.: 3,77e-120 Length: 510
Score: 1834.00 Matches: 364
Percent Similarity: 78.57% Conservative: 43
Best Local Similarity: 70.27% Mismatches: 71
Query Match: 68.15% Indels: 40
DB: 2 Gaps: 8

US-10-728-051-3 (1-1524) x Q61WG5 (1-510)
QY 1 CGGACGACACCGGAGAGAGACCGGTGCCAGTTCAGAGCGCTCAATGGCGAGAGACTGAC 60
Db 22 ArgGlnGlyGlyGluGluAsnGluCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 41
QY 61 AATCCGATTTGAATCAGAGCGGTTCATTTGAGACTTGGAAACCCCAACCAACAGGAGTTC 120
Db 42 AsnArgIleGluSerGlnGlyGlyTyrIleGluThrTrpAsnProAsnGlnGluPhe 61
QY 121 GAATCGCGCGGTGCGCCCTCTCTCGTTCGTAGTCTCCGCGCAACGCCCTTCGTAGGCT 180
Db 62 GlnCysAlaGlyValAlaLeuSerArgThrValLeuArgAsnAlaLeuArgPro 81
QY 181 TTCTACTCTCAATGCTCCCGAGAGATCTTCATCCAGCAGAGGAGGGGATCTTTGGGTG 240
Db 82 PheTyrSerAsnAlaProLeuGluIleTyrValGlnGlnGlySerGlyTyrPheGlyLeu 101
QY 241 ATATTCCCTGGTGTGCTTAGACATATCAAGAGCTTCACACACAGGTCTCGATCTCTCAG 300
Db 102 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGluGlyArgArgTyrGln 121
QY 301 TCCCAAGACCAACCAAGACGTCTCCAA-----GGAGAAGACCAAGCCCAAGCAACGA 354
```

```
Db 122 SerGlnLysProSerArgArgPheGlnValGlyGlnAspAspProSerGlnGlnGlnGln 141
QY 355 GATPAGTCACCAAGAGTGCCCGTTCGATGAGGGTGCATCTCATTCGAGTTCACCGCGT 414
Db 142 AspSerHisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGly 161
QY 415 GTTGTCTTCTGCTCTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 474
Db 162 ValAlaPheThrPheMetTyrAsnAspGluAspThrAspValValThrValThrLeuSerAsp 181
QY 475 ACCAAACAACACACACACACACACACACACACACACACACACACACACACACACAC 534
Db 182 ThrSerSerIleHisGlnLysGlnLeuAspGlnPheProArgArgPheTyrLeuAlaGlyAsn 201
QY 535 ACCGAGCAGAGGTTCTTAAGGTACACGACCAACAAAGCAGACAAAGCAGACGACGAAGCTTA 594
Db 202 GlnGluGlnGluPheLeuArgTyrGlnGlnGlnGlySer----- 215
QY 595 CCATATAGCCCATACAGCCCGCAAGTGCACCTAGACAAAGAGCGTGATTTAGCCCT 654
Db 216 -----ArgProHisTyrArgGlnIleSerPro 224
QY 655 CGAGGACAGCACACGCCGCGCAGAGAACGAGCAGACGAGCAAGAAACGAGGTGGAAC 714
Db 225 ArgValArg-----GlyAspGlnGlnGlnAsnGluGlySerAsn 237
QY 715 ATCTTCAGCGGCTTCAGCGCGAGTTCCTGGAAACAAGCCCTTCAGGTTGACGACAGACAG 774
Db 238 IlePheSerGlyPheAlaGlnGluPheLeuGlnHisAlaPheGlnVal---AspArgGln 256
QY 775 ATAGTCAAAACCTAGAGCGGAGACCGAGAGTGAAGAGAGGAGGAGCATTGTGACAGTG 834
Db 257 ThrValGluAsnLeuArgGlyGluAsnGluArgGluGlnGlyAlaIleValThrVal 276
QY 835 AGGGAGGCTCAGATCTTCAGCCAGCAGATAGAAAGAGAGCGTGCC----- 879
Db 277 LysGlyGlyLeuArgIleLeuSerProAspGluAspGluSerArgSerArgSerProPro 296
QY 880 GACGAAGAAGAGGAATACGATGAAGATGAATATATACGATGAAGAGGATGAAGCGGT 939
Db 297 SerArgArgGluGluPheAspGluAspArg-----SerArgProGlnGlnArg 312
QY 940 GGCAGG-----GGAACGACAGGCGAGGGGAAATGGTATTGAAGAGACGATCTGCACC 990
Db 313 GlyLysTyrAspGluAsnArgArgGlyTyrLysAsnGlyIleGluGluThrIleCysSer 332
QY 991 GCAAGTGCTAAAGACATCTGCTAGAACACAGATCCCTGACATCTACACCCCTCAAGCT 1050
Db 333 AlaSerValLysLysAsnLeuGlyArgSerSerAsnProAspIleTyrAsnProGlnAla 352
QY 1051 GGTTCACCTCAAACTGCCAACGATCTCAACCTTCTAATCTAGTGGCTTGACCTAGT 1110
Db 353 GlySerLeuArgSerValAsnGluLeuAspLeuProIleLeuGlyTrpLeuGlyLeuSer 372
QY 1111 GCTGAATATGAATCTCTACAGAAATGCATTTGCTGCTCCTCACTACACACCAACGCA 1170
Db 373 AlaGlnHisGlyThrIleTyrArgAsnAlaMetPheValProHisTyrThrLeuAsnAla 392
QY 1171 CACAGCATCATATCGATTGAGGGGAGCGGCTCAGTGCAGTGCAGTGCAGCAGCAGCGC 1230
Db 393 HisThrIleValValAlaLeuAsnGlyArgAlaHisValGlnValValAspSerAsnGly 412
QY 1231 AACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACGTGCTGTGGTGCACAGAACTTC 1290
Db 413 AsnArgValTyrAspGluGluLeuGlnGluGlyHisValLeuValValProGlnAsnPhe 432
QY 1291 GCCGTGCTGGAAGTCCCAAGAGCGAAGTTCGAATACGTGGCATTCACAGACAGACTCA 1350
Db 433 AlaValAlaAlaLysAlaGlnSerGluAsnTyrGluTyrLeuAlaPheLysThrAspSer 452
QY 1351 AGGCCAGCATAGCCCACTCGCGGTGAAACCTCCGTCTATAGATAACCTCGCGGAGGAG 1410
```

```
Db 453 ArgProSerIleAlaAsnLeuAlaGlyGluAsnSerIleIleAspAsnLeuProGluGlu 472
QY 1411 GTGGTTGCAAAATTCATATGCGCTCCAAAGGAGCAGCAGGAGGAGGAGGAGGAGGAG 1470
Db 473 ValValAlaAsnSerTyrArgLeuProArgGluGlnAlaArgGlnLeuLysAsnAsnAsn 492
QY 1471 CCCTTCAAGTTCTTCGTTCCACCG---TCTCAGCAGTCTCCGAGGCGCTGTGGCT 1521
Db 493 ProPheLysPhePheValProPheAspHisGlnSerMetArgGluValAla 510

RESULT 9
Q647H2 PRELIMINARY; PRT; 484 AA.
ID Q647H2;
AC Q647H2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Arachin Ahv-3.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV722687; AAU21492.1; -.
SQ SEQUENCE 484 AA; 54568 MW; 5A3E950752E89D2D CRC64;

Alignment Scores:
Pred. No.: 2,58e-107 Length: 484
Score: 1651.00 Matches: 337
Percent Similarity: 75.25% Conservative: 46
Best Local Similarity: 66.21% Mismatches: 76
Query Match: 61.35% Indels: 50
DB: 2 Gaps: 8

US-10-728-051-3 (1-1524) x Q647H2 (1-484)
QY 1 CGGCAGCAACCGGAGAGAACCGGTCGCTCCAGGCTCAATCGCGCAGACCTGAC 60
Db 24 ArgGlnGlnGlyGluGluAsnGluCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCGCATTTGAATTCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACACAGGAGTTC 120
Db 44 AsnCysIleGluSerGluGlyIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
QY 121 GAATGCGCGGCGTCCGCTCTCTCGCTTAGTCTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 64 GlnCysAlaGlyValAlaLeuSerArgPheValLeuArgAsnAlaLeuArgPro 83
QY 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGGAGTTCCTTGGGTG 240
Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleTyrGlnGlySerGlyTyrPheGlyLeu 103
QY 241 ATATTCCTCGTGTGCTTAGACACTATGAAGAGCCTCACACAAAGCTCGTCGATCTCAG 300
Db 104 IlePheProGlyCysProGlyThrPheGluGluPro---IleGlnGly---SerGluGln 121
QY 301 TCCCAAGACCAACAGACGCTCTCCAAAGGAGAGGACCAAGCCCAACAGCAGAGTAGT 360
Db 122 PheGlnArgProSerArgHisPheGlnGlyGlnAspGlnSerGlnArgProLeuAspThr 141
QY 361 CACGAGAGGTGACCGGTTTCGATGAGGTGATCTCATTCGAGTTCCTCCACCGGTGTGCT 420
Db 142 HisGlnLysValHisGlyPheArgGluGlyAspLeuIleAlaValProHisGlyValAla 161
QY 421 TTCTGCTCTACACCAACGACGACACTGATGTTGTTGTTGTTCTCTTACTGACACCAAC 480
Db 162 PheThrIleTyrAsnAspGlnAspThrAspValValAlaIleSerValLeuHisThrAsn 181
```

```
QY 481 AACACGACACACGAGTGTGATGATGTTCCCCAGGAGATTCATTTGGCTGGGAACACGGAG 540
  :: |||||
Db 182 SerLeuHisAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyLysGlnGlu 201
  |||||
QY 541 CAAGAGTCTTAAGTACACGACCAACAGC---AGACAAAGCAGCAGCAAGCACTTACCA 597
  |||||
Db 202 GlnGluPheLeuArgfyrGlnGlnArgSerGlyArgGlnSer 215
  |||||
QY 598 TATAGCCCATACAGCCCGCAAGTGCAGCTAGACAAAGAGAGCGTGAATTTAGCCCTCGA 657
  |||||
Db 216 -----ProlyseGlyGluGlnGlu----- 222
QY 658 GGACAGCACGCCGACAGAACAGCAGGACCAAGAAACCAAGAGTGGAAACATC 717
  |||||
Db 223 -----GlnGluGlnGluAsnGluGlyGlyAsnVal 232
QY 718 TTCAGCGGCTTACGCCGAGTTCCTGGAACAGCCTTCAGGTTACGACGACAGATA 777
  |||||
Db 233 PheSerGlyPheSerThrGluPheLeuSerHisGlyPheGlnValAsnGlu---AspIle 251
  |||||
QY 778 GTCAAAACCTAAGAGCGCAGACCGCAGAGTGAAGAGAGAGGAGGCCATTGTGACAGTGAGG 837
  |||||
Db 252 ValArgAsnLeuArgGlyGluAsnGluArgGluGlnGlnGlyAlaIleValThrValLys 271
  |||||
QY 838 GGAGGCTCAGATCTTGAGCCCGACAGATAGAGAGAGAGAGAGAGAGAGAGAGAGATAC 897
  |||||
Db 272 GlyGlyLeuSerIleLeuValPro----- 280
QY 898 GATGAAGTGAATATGAATACGATGATGAGGATAGAGGCGTGGCAGGAGGACGAGGC 957
  |||||
Db 281 -----GluTrpArgGlnSerTyrGlnGlnProGlyArgGlyAspLysAsp 295
QY 958 AGGGGAATGTTGAAGAGAGCGATCTGCACCGCAAGTGCTAAAGAACATTTGTGAGA 1017
  |||||
Db 296 PheAsnAsnGlyIleGluGluThrIleCysThrAlaThrValLysMetAsnIleGlyLys 315
QY 1018 AACAGATCCCTGACATCTACACCTCAAGCTGTGTTCTCAAACTGCCAACGATCTC 1077
  :: |||||
Db 316 SerThrAlaAspIleTyrAsnProGlnAlaGlySerValArgThrValAsnGluLeu 335
QY 1078 AACCTTCTAATCTAGTGGCTTGCACCTAGCTGCTGAATATGGAATCTCTACAGGAAT 1137
  :: |||||
Db 336 AspLeuProIleLeuAsnArgLeuGlyLeuSerAlaGluTyrGlySerIleHisArgAsp 355
QY 1138 GCATTTGTTGCTCACTACACCAACGACGACGACATCATATCATTCATTGAGGGA 1197
  :: |||||
Db 356 AlaMetPheValProHisTyrAsnMetAsnAlaAsnSerMetIleTyrAlaLeuHisGly 375
QY 1198 CGGGCTCAGTGAAGTGTGGACAGCAACGCGCAACAGAGTGTACACGAGGAGCTTCAA 1257
  |||||
Db 376 GlyAlaHisValGlnValValAspCysAsnGlyAsnArgValPheAspGluGluLeuGln 395
QY 1258 GAGGTCACGTGCTGTGTGTGCACAGAACTTCGCGTGCCTGGAAGTCCCGAGCGAG 1317
  |||||
Db 396 GluGlyGlnSerLeuValValProGlnAsnPheAlaValAlaAlaLysSerGlnSerGlu 415
QY 1318 AACTTCGAATACGTGGCTTCAAGACAGACTCAAGGCCCGCAGCATACCACTCCGCGGT 1377
  :: |||||
Db 416 HisPheLeuTyrValAlaPheLysThrAsnSerArgAlaSerIleSerAsnLeuAlaGly 435
QY 1378 GAAACTCCGTATAGATAACCTGCGGAGGAGGTGTGTCAAATTCATATGCTCCCAA 1437
  :: |||||
Db 436 LysAsnSerTyrMetTrpAsnLeuProGluAspValValAlaAsnSerTyrGlyLeuGln 455
QY 1438 AGGAGCAGCAGGAGCTTGAAGAACAAACACCCCTTCAAGTTCCTGCTCCACG--- 1494
  |||||
Db 456 TyrGluGlnAlaArgGlnLeuLysAsnAsnProPheThrPheLeuValProGln 475
QY 1495 TCTACAGCTCTCCGAGGCGTGTGCT 1521
  |||||
Db 476 AspSerGlnMetIleArgThrValAla 484
```

## RESULT 10

```
GLC3_SOYBN STANDARD; PRT; 481 AA.
AC P11828;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Glycinin G3 precursor [Contains: Glycinin A subunit; Glycinin B
  subunit].
GN Name-GY3;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]_TaxID=3847;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=89296500; PubMed=2740231;
RA Cho T.-J., Nielsen N.C.;
RT "The glycinin G3 gene from soybean.";
RL Nucleic Acids Res. 17:4388-4388(1989).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=9239391; PubMed=2485233;
RA Nielsen N.C., Dickinson C.D., Cho T.-J., Thanh V.H., Scallan B.J.,
RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;
RT "Characterization of the glycinin gene family in soybean.";
RL Plant Cell 1:313-328(1989).
CC -!- FUNCTION: Glycinin is the major seed storage protein of soybean.
CC -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond.
CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
  family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
DR EMBL; X15123; CAA33217.1; -.
DR PIR; S04605; S04605.
DR HSP; P04776; 1EXZ.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmlC like cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S SEED STORAGE; 1.
KW Multigene family; Seed storage protein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 296 Glycinin A subunit.
FT CHAIN 297 476 Glycinin B subunit.
FT PROPEP 477 481
FT DISULFID 107 303 Interchain (By similarity).
SQ SEQUENCE 481 AA; 54241 MW; 5F3C3148DF6241A7 CRC64;
Alignment Scores:
Pred. No.: 7,33e-97 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.57% Conservative: 68
Best Local Similarity: 58.24% Mismatches: 91
Query Match: 1 Indels: 54
DB: 1 Gaps: 7
US-10-728-051-3 (1-1524) x GLC3_SOYBN (1-481)
QY 1 CGGACGACACCGGAGGAGAAACCGTGCACGCTTCAGCGCTTCATGCGAGAGACCTGAC 60
```

Db	23	ArgGlnProGlnAenGluCyGlnIleGlnArgLeuAsnAlaLeuIysProAsp	42
Qy	61	AATCGCATTAATCAGAGGCGGTACATTGACAGCTTGGAAACCCCAACACCAAGGAGTTTC	120
Db	43	AsnArgIleGluSerGluGlyGlyPheIleGluThrIrpAsnProAsnAsnIysProPhe	62
Qy	121	GAATGCGCCGCGCTCGCCCTCTCTCGCTTAGTCTCTCCGCGCAACCCCTTCGTAGGCTT	180
Db	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro	82
Qy	181	TTCTACTCCAATGTCCTCCAGGAGATCTTCATCCAGCAAGGAGGGATACITTTGGTTG	240
Db	83	SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet	102
Qy	241	ATATTCCCTGTTGCTCTAGACACTATGAAGAGCCTCACACACAAGTCGTGCATCTCAG	300
Db	103	IlePheProGlyCysProSerThrPheGluGluProGlnGlnIysGly	119
Qy	301	TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAAGCCAAACCAACGACCAACGAGATAGT	360
Db	120	SerSerArgPro	126
Qy	361	CACCAAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCACCGGTGTTGCT	420
Db	127	HisGlnIysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla	146
Qy	421	TTCTGGCTCTACACGACCAACGACACGTGTTGTTGCTGTTCTCTTACTGACACCAAC	480
Db	147	TyrIrpMetTyrAsnAenGluAepThrProValValAlaValSerLeuIleAspThrAsn	166
Qy	481	AACAACGACCAACCAAGCTTGATCAGTTCCCGAGGATTCATTTGGCTGGGAACACGGAG	540
Db	167	SerPheGlnAenGlnLeuAepGlnMetProArgPheTyrLeuAlaGlyAenGlnGlu	186
Qy	541	CAAGAGTTCTTAAGGTACAGCAACAAGACAGACAAGACAGACAAGAGCTTACCATAT	600
Db	187	GlnGluPheLeuGlnTyrGlnProGlnIysGlnGlnGly	199
Qy	601	AGCCCATACAGCCCGCAAGCTCAGCTAGACAAGAAGCGTGAAATTAGCCCTCGAGGA	660
Db	200		200
Qy	661	CAGCACAGCCGACAGAACGACGACGAGCAAGAAGAAACCAAGGTGGAAACATCTTC	720
Db	201	ThrGlnSerGlnIysGlyLysArgGlnGlnGluGluAenGluGlySerIleLeu	220
Qy	721	AGCGGCTTCAGCCGGAGTTCCTGGACAACGCTTCAGGTGACGACAGACAGATAGTG	780
Db	221	SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal	239
Qy	781	CAAACTTACAGGCGGACCGACAGTGAAGAGAGGAGGCCATTTGTGACAGTAGGGGA	840
Db	240	ArgIysLeuGlnGlyGluAenGluGluGluIysGlyAlaIleValThrValIysGly	259
Qy	841	GGCCTCAGAACTTCTGAGCCCA	891
Db	260	GlyLeuSerValIleSerProProThrGluGluGlnGlnArgProGluGluGlu	279
Qy	892	GAATACGATGAAGATGAATATGAATATCATGATGAAGAGATGAAGGCGTCGCGAGGAGC	951
Db	280	LysProAsp	294
Qy	952	AGAGGCGGGGGAATGTTATTGAAGAGACGATCTGCACCGCAAGTCGTCTAAAAGACATT	1011
Db	295	Arg	311
Qy	1012	GGTAGAACAGATCCCTTGACATCTACAACCCCTCAAGCTGGTTCACATCAAACTGCCAAC	1071
Db	312	GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr	331
Qy	1072	GATCTCAACCTTCTAACTTAGTGGCTTGGACCTAGTGTGCTGAATATGAAATCTCTAC	1131

Db	332	SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg	351
Qy	1132	AGGAATGCAATGTTGTTGTCGCTCACTACAACACCAACGACACAGCATCATATATCGATTG	1191
Db	352	LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu	371
Qy	1192	AGGGACGGGCTCACGTGCAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAG	1251
Db	372	AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu	391
Qy	1252	CTTCAAGAGGGTCACGTGCTTGTTGGTSCCACAGAACTTCGCCGTGCGTGGAAGTCCCAG	1311
Db	392	LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln	411
Qy	1312	AGCGAGAACTTCGAATACGTGGCATTCACAGACAGACTCAAGGGCCAGCATAGCAACCTC	1371
Db	412	SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu	431
Qy	1372	GCGGTCGAAACTCCGTCATAGATAACCTGCCGGAGGAGGTGGTTGCAAAATTCATATGGC	1431
Db	432	AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn	451
Qy	1432	CTCCAAAGGAGGACGACGCAAGCGAGCTTAAGCAACAAACCCCTCAAGTTCTTTCGTTCCA	1491
Db	452	LeuArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro	471
Qy	1492	CCGTCTCAGCAGTCTCCGAGGGCTGTGGCT	1521
Db	472	ProLysGluSerGlnArgValValAla	481
RESULT 11			
GLC2 SOYBN			
ID	GLC2 SOYBN	STANDARD;	PRT; 485 AA.
AC	P04405; P04121; P04348; P04349;		
DT	20-MAR-1987 (Rel. 04, Created)		
DT	01-OCT-1989 (Rel. 12, Last sequence update)		
DT	05-JULI-2004 (Rel. 44, Last annotation update)		
DE	Glycinin G2 precursor [Contains: Glycinin A2 subunit; Glycinin B1a subunit]		
GN	Name=GV2;		
OS	Glycine max (Soybean).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.		
OX	NCBI_TaxID=3847;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Dare; TISSUE=Leaf;		
RX	MEDLINE=92393391; PubMed=2485233;		
RA	Nielsen N.C., Dickinson C.D., Cho T.J., Thanh V.H., Scallan B.J.,		
RA	Fischer R.L., Sams T.L., Drews G.N., Goldberg R.B.;		
RT	"Characterization of the glycinin gene family in soybean.";		
RL	plant Cell 1:313-328(1989).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Dare; TISSUE=Leaf;		
RX	MEDLINE=89296499; PubMed=2740230;		
RA	Thanh V.H., Tumer N.E., Nielsen N.C.;		
RT	"The glycinin Gy2 gene from soybean.";		
RL	Nucleic Acids Res. 17:4387-4387(1989).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Shiotsurunoko;		
RA	Utsumi S., Kim C.S., Kohno M., Kito M.;		
RT	"Polymorphism and expression of cDNAs encoding glycinin subunits.";		
RL	Agric. Biol. Chem. 51:3267-3273(1987).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=88040439; PubMed=3671077;		
RA	Fukazawa C., Momma T., Higuchi W., Uda K.;		
RT	"Complete nucleotide sequence of the gene encoding a glycinin A2B1a subunit precursor of soybean.";		
RL	Nucleic Acids Res. 15:8117-8117(1987).		







Db 238 ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 257  
QY 841 GGCTCAGAACTTGAGCCAGATAGAAAGAGAGCGTCCGACGAGAGAGGAAATACGAT 900  
Db MEDLINE=89296498; PubMed=2740229;  
258 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAsp 277  
QY 901 GAAGATGAA-----TATGATACATGACAGAGATAGAGGCTGCAGGGGA 948  
Db 278 GluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 297  
QY 949 AGCAGAGCGAGGGGAATGATTAAGAGACCATCTGCACCCCAAGTGTCTAAAGAAAC 1008  
Db 298 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 314  
QY 1009 ATTGGTAGAAACAGATCCCTGATCATCTACAACCTCAAGCTGTTCACTCAAAATGCC 1068  
Db 315 IleGlyGlnAsnSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 334  
QY 1069 AACGATCTCAACCTTCTAATACTAGTGGCTTGGACCTAGTCTGAATATGGAATCTC 1128  
Db 335 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 354  
QY 1129 TACAGGAATCATTTGTTGCTGCTACTACAACACCAACGACACAGCATCATATATCGA 1188  
Db 355 ArgIysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 374  
QY 1189 TTGAGGGGACGGCTCAGTGCAGTCTGCGAGCAGCAACGCGACAGAGTGTACGACGAG 1248  
Db 375 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 394  
QY 1249 GAGCTTCAAGAGGTACGTGCTTGTGTCGTCAGCAACTTCGCGCTCGTGGAAAGTCC 1308  
Db 395 GluLeuGlnGlyGlyValLeuIleValProGlnAsnPheAlaValAlaLysSer 414  
QY 1309 CAGAGGAGAACTTCGAATACGTGGCATTCAGACAGCTCAAGCCACATAGCAAC 1368  
Db 415 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434  
QY 1369 CTCGCCGCTGAAACTCCGTATAGATAAATCTGCGGAGGAGGTGGTTCGAAATTCAT 1428  
Db 435 LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluValIleGlnHisThrPhe 454  
QY 1429 GGCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1488  
Db 455 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 474  
QY 1489 CCACCGCTCAGCAGTCTCCGAGGCGTGTGCT 1521  
Db 475 ProProGlnGluSerGlnArgArgAlaValAla 485

RESULT 12  
GLC1\_SOYBN STANDARD; PRT; 495 AA.  
AC P04776;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Glycinin G1 precursor [Contains: Glycinin A1A subunit; Glycinin B1 subunit].  
DE Name=Gly;  
GN Glycine max (Soybean).  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
ON NCBI\_TaxID=3847;  
RX [1]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Dare; TISSUE=Leaf;  
RX MEDLINE=92393391; PubMed=2485233;  
RA Nielsen N.C., Dickinson C.D., Cho T.J., Thanh V.H., Scallion B.J.,  
RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;  
RT "Characterization of the glycinin gene family in soybean."  
RL Plant Cell 1:313-328(1989).

[2] SEQUENCE FROM N.A.  
RC STRAIN=cv. Dare; TISSUE=Leaf;  
RX MEDLINE=89296498; PubMed=2740229;  
RA Sims T.L., Goldberg R.B.;  
RT "The glycinin Gyl gene from soybean."  
RL Nucleic Acids Res. 17:4386-4386(1989).  
[3] SEQUENCE FROM N.A.  
RC STRAIN=cv. Bonminori;  
RX MEDLINE=86041867; PubMed=2997720;  
RA Negoro T., Momma T., Fukazawa C.;  
RT "A cDNA clone encoding a glycinin A1A subunit precursor of soybean."  
RL Nucleic Acids Res. 13:6719-6731(1985).  
[4] SEQUENCE FROM N.A.  
RC STRAIN=cv. Kito M.;  
RA Utsumi S., Kohno M., Mori T., Kito M.;  
RT "An alternate cDNA encoding glycinin A1A Bx subunit."  
RL J. Agric. Food Chem. 35:210-214(1987).  
[5] SEQUENCE FROM N.A.  
RC Urade R., Nakatani H., Nakano C.;  
RT "mRNA of soybean proglycinin A1A1B1 subunit."  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
CC -I- FUNCTION: Glycinin is the major seed storage protein of soybean.  
CC -I- SUBUNIT: Hexamer; each subunit is composed of an acidic and a  
CC disulfide bond.  
CC -I- PTM: The precursor is post-translational processed to form a  
CC covalently linked A1A-BX subunit complex.  
CC -I- SIMILARITY: Belongs to the 11S seed storage protein (globulins)  
CC family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; M36686; AAA33966.1; -  
CC EMBL; X15121; CAA33215.1; -  
CC EMBL; X02985; CAA36723.1; -  
CC EMBL; AB113349; BAC78522.1; -  
CC PIR; A23497; FMSY62.  
CC PIR; S10851; S10851.  
CC PDB; 1FXZ; X-ray; A/B/C=20-495.  
CC PDB; 1UCX; X-ray; A/B/C=20-495.  
CC PDB; 1UD1; X-ray; A/B/C=20-495.  
CC InterPro; IPR006045; Cupin.  
CC InterPro; IPR007113; Cupin region.  
CC InterPro; IPR011051; RmlC like cupin.  
CC InterPro; IPR006044; Seedstore\_11s.  
CC Pfam; PF00190; Cupin; 2.  
CC PRINTS; PR00439; 11SGLOBULIN.  
CC PROSITE; PS00305; 11S SEED STORAGE; 1.  
KW 3D-structure; Multigene family; Seed storage protein; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 306 Glycinin A1A subunit.  
FT PROPEP 307 310  
FT CHAIN 311 490 Glycinin B1 subunit.  
FT PROPEP 491 495  
FT DISULFID 107 317 Interchain (By similarity).  
FT CONFLICT 42 42 D -> G (in Ref. 3).  
FT CONFLICT 108 108 P -> S (in Ref. 3).  
FT CONFLICT 136 136 E -> S (in Ref. 3).  
FT CONFLICT 360 360 E -> G (in Ref. 3).  
FT TURN 30 31  
FT STRAND 39 39  
FT STRAND 43 47  
FT TURN 48 49  
FT STRAND 50 54

FT	TURN	57	58		
FT	HELIX	60	65		
FT	TURN	66	66		
FT	STRAND	68	74		
FT	TURN	76	77		
FT	STRAND	78	84		
FT	STRAND	89	95		
FT	STRAND	97	103		
FT	TURN	105	106		
FT	STRAND	134	137		
FT	TURN	138	139		
FT	STRAND	140	144		
FT	TURN	146	147		
FT	STRAND	149	154		
FT	STRAND	160	166		
FT	TURN	168	169		
FT	TURN	171	172		
FT	STRAND	181	183		
FT	TURN	191	192		
FT	HELIX	193	196		
FT	TURN	222	223		
FT	HELIX	226	233		
FT	TURN	234	234		
FT	HELIX	237	243		
FT	STRAND	255	257		
FT	TURN	317	318		
FT	STRAND	322	324		
FT	STRAND	333	335		
FT	TURN	336	338		
FT	STRAND	339	344		
FT	TURN	346	348		
FT	HELIX	350	352		
FT	TURN	353	356		
FT	STRAND	359	364		
FT	TURN	366	367		
FT	STRAND	369	375		
FT	TURN	376	377		
FT	STRAND	380	386		
FT	STRAND	388	394		
FT	TURN	396	397		
FT	STRAND	400	407		
FT	TURN	408	409		
FT	STRAND	411	414		
FT	TURN	416	417		
FT	STRAND	419	424		
FT	TURN	427	428		
FT	STRAND	429	435		
FT	STRAND	442	444		
FT	TURN	448	449		
FT	TURN	451	451		
FT	HELIX	452	454		
FT	HELIX	457	464		
FT	TURN	465	465		
FT	HELIX	468	476		
FT	STRAND	483	484		
SQ	SEQUENCE	495 AA;	55706 MW; 329CB0545B24D894 CRC64;		

Alignment Scores:  
Pred. No.: 1 94e-96 Length: 495  
Score: 1496.00 Matches: 292  
Percent Similarity: 70.87% Conservative: 73  
Best Local Similarity: 56.70% Mismatches: 100  
Query Match: 55.59% Indels: 50  
DB: 1 Gaps: 6

US-10-728-051-3 (1-1524) x GLC1\_SOYBN (1-495)

QY	1	CGCAGCAACCGGAGGAGAACGGTCCAGCTCCAGCGCTCAATCGCAGAGACCTGAC	60
Db	23	ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp	42
QY	61	AATCGCATTAATCAGAGGGCGGTACATTGACACTTGGNACCCCAACACAGGAGTTTC	120

Db	43	AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysProPhe	62
QY	121	GAATCGCGCGGCTGCGCTCTCTCGCTTAGCTCCGCGCAAGCCCTTCGTAGGCCT	180
Db	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro	82
QY	181	TTCTACTCCAATGCTCCCGCAGAGATCTTCATCCAGCAAGGAAGGATACCTTTGGGTG	240
Db	83	SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysGlyIlePheGlyMet	102
QY	241	ATATTCCCTGCTGCTAGACACATATGAAGAGCTTCACACAAAGGTGCTCGATCTCAG	300
Db	103	IleTyrProGlyCysProSerThrPheGluGluProGlnGlnProGlnArgGlyGln	122
QY	301	TCCCAAGACACCAAGACGCTCTCCAAGGAGAAGACCAAGCCAAAGCAGCAACGAGATAGT	360
Db	123	SerSerArgPro-----GlnAspArg	129
QY	361	CACCAGAAAGGTGCACCGTTTCGATGAGGTGATCTATTGCAGTTCACACCGGTGTGCT	420
Db	130	HisGlnLysIleTyrAsnPheArgGluGlyAspLeuIleAlaValProThrGlyValAla	149
QY	421	TTCTGCTCTACAAACGACACGACACTGATGTGTGCTTCTTCTTACTCACCACAC	480
Db	150	TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleAspThrAsn	169
QY	481	AACAACGACAAACAGCTTGATCAGTTCCCGCAGAGATTCAATTTGGCTGGGAACACGGAG	540
Db	170	SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	189
QY	541	CAAGAGTTCTTAAGGTACCAACAAAGCAGACAAAGCAGAGCAAGAGCTTACCATAT	600
Db	190	GlnGluPheLeuLysTyrGlnGlnGln-----	199
QY	601	AGCCCATACAGCCCGCAAGTCCAGCTAGACAGAGCGTGAATTTAGCCCTCGAGCA	660
Db	200	-----GlyGly	201
QY	661	CAGCACAGCCGCGAGAGACGAGCAGACAAGAAAGAAAGAAAGAGTGGAAACATCTTC	720
Db	202	HisGlnSerGlnLysGlyHisGlnGlnGlnGluGluAsnGlnGlySerIleLeu	221
QY	721	AGCGCTTCACCGCGAGTTCTCTGGAACAAAGCTTCCAGTTGACGACAGACATAGTG	780
Db	222	SerGlyPheThrLeuGluPheLeuGluHisAlaPheSerVal---AspLysGlnIleAla	240
QY	781	CAAAACCTAAGAGCGCAGACCGAGAGTGAAGAGGAGGAGCCATTGTGACGTGAGGGA	840
Db	241	LysAsnLeuGlnGlyGluAsnGluGlyGluAspLysGlyAlaIleValThrValLysGly	260
QY	841	GGCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGACGTCCCGCAGCAAGAGAG	891
Db	261	GlyLeuSerValIleLysProThrAspGluGlnGlnGlnArgProGlnGluGlu	280
QY	892	GAATACGATGAAGATGAATATGAATAC-----GATGAAGAG---GATGAAGAG	936
Db	281	GluGluGluGluAspLysProGlnCysLysGlyLysAspLysHisCysGlnArgPro	300
QY	937	CGTGGCAGGGGAGCAGAGCGGCGGAGTGTATTGAAGACGATCTGCACCGCAAT	996
Db	301	ArgGlySerGlnSerLysSerArgArgAsnGlyIleAspGluThrIleCysThrMetArg	320
QY	997	GCTAAAAGAAACATTGGTAGAAACAGATCCCTGACATCTACAACCTCAAGCTGGTTCA	1056
Db	321	LeuArgHisAsnIleGlyGlnThrSerSerProAspIleTyrAsnProGlnAlaGlySer	340
QY	1057	CTCAAAATCGCAACGATCTCAACCTTCTAATPACTTAGTGGCTGGACCTAGTCTGAA	1116
Db	341	ValThrThrAlaThrSerLeuAspPheProAlaLeuSerTrpLeuArgLeuSerAlaGlu	360
QY	1117	TATGGAAATCTCTACAGGAATCATTGTTGCTGCTACTACAACACCAACCAACGACAC	1176
Db	361	PheGlySerLeuArgLysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSer	380

Db	23	ArgGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProGly	42
Qy	61	AATGCGATTGAATCAGAGGCGGTATCATTTGAGACTTGGAACCCCAACACACGAGAGTTC	120
Db	43	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe	62
Qy	121	GAATGGCGGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGGCCT	180
Db	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgA-gpro	82
Qy	181	TTCTACTCCAATGCTCCCAAGGAGATCTTCATCCAGCAAGAAAGGGGATACTTTGGGTG	240
Db	83	SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet	102
Qy	241	ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCCTCACACACAAGTTCGTCACTTCAG	300
Db	103	IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly	119
Qy	301	TCCCAAGACCACCAAGACGTCTCCAAAGGAGAGACCAAGCCACACGCAACAGAGATAGT	360
Db	120	SerSerArgPro	126
Qy	361	CACCAGAAGTGCACCGTTTCGATGAGGGTGATCTCATTCGAGTTCACCGGTGTGTCT	420
Db	127	HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleValProThrGlyPheAla	146
Qy	421	TTCTGGCTCTAACACGACCAACGACACTGATGTTGTGCTGTTCTCTTACTGCACCAAC	480
Db	147	TyrTrpMetTyrAsnAsnGluAspThrLeuValValAlaValSerLeuIleAspThrAsn	166
Qy	481	AACACGACCAACGAGCTTGATCAGTTCCTCCGAGAGATTCAATTTGGCTGGGAACACGGAG	540
Db	167	SerPheGlnAsnGlnLeuAspGlnMetProArgPheTyrLeuAlaGlyAsnGlnGlu	186
Qy	541	CAAGAGTCTTAAGTTACCAACCAACAAAGCAGACAAAGCAGACGACGAAGACCTTACCATAT	600
Db	187	GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly	199
Qy	601	AGCCCATACAGCCGCAAAAGTCAGCCTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGGA	660
Db	200	-----Gly	200
Qy	661	CACGACAGCGCAGAGAACGAGCAGGACAGAGAGAGAAACGAAAGTGGAACATCTTC	720
Db	201	ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlyGlySerMetLeu	220
Qy	721	AGCGGCTTACGCGCGAGTTCCTGGACAAGCCCTCCAGGTTCCAGCACGACGACGATAGTC	780
Db	221	SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal	239
Qy	781	CAAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGAAGGGA	840
Db	240	ArgLysLeuGlnGlyGluAsnGluGluGluLysGlyAlaIleValThrValLysGly	259
Qy	841	GGCCTCAGAATCTTGAGCCCA-----GATGAAAGAGACGTCGCCGACGAAGAGAG	891
Db	260	GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu	279
Qy	892	GAATACGATCAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGCACGGGAGAC	951
Db	280	LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer	294
Qy	952	AGAGGCGAGGGGAATGTTATTGAAGAGACGATCTGCACCGCAAGTGCTAAAAAGAACATT	1011
Db	295	Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle	311
Qy	1012	GGTAGAAACAGATCCCTCGACATCTACAAACCTCAAGCTGGTTCATCTCAAACTGCCAAC	1071
Db	312	GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr	331
Qy	1072	GATCTCAACCTTCTAATACCTTAGGTGGCTGGACCTAGTGCTGAATGATGAATCTCTAC	1131
Db	332	SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg	351

**Figure 1**



```

Db 353 LyasAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 372
QY 1192 AGGGAGCGGCTCAGCTGCAAGTCGTGGAGCAGCAACGCGCAACAGAGTGTCAGCAGGAG 1251
Db 373 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 392
QY 1252 CTTCAAGAGGGTCAGCTGCTGTGGTGGCCACAGAACTTCGCCCTCGCTGGAAAGTCCCAG 1311
Db 393 LeuGlnGluGlyGlnValLeuThrValProGlnAsnPheAlaValAlaAlaArgSerGln 412
QY 1312 AGCGAGAACTCCGAATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCAACCTC 1371
Db 413 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 432
QY 1372 GCCGGTGAAACTCCGCTAGATTAACCTCCCGAGGAGGTGTGCAAAATTCATATGGC 1431
Db 433 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 452
QY 1432 CTCGAAGGAGGAGGAGGAGGAGCTTAAGAAACAACCCCTTCAGTTCTTCGTTCGA 1491
Db 453 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnProPheSerPheLeuValPro 472
QY 1492 CCGTCTCAGAGTCTCCGAGGGCTGTGGCT 1521
Db 473 ProLysGluSerGlnArgArgValAla 482

RESULT 15
Q41702
ID Q41702 PRELIMINARY; PRT; 498 AA.
AC Q41702;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Legumin A precursor.
OS Vicia sativa (Spring vetch) (Tare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OX NCBI_TaxID=3908;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cotyledon;
RA Nong V., Becker C., Muentz K.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond (By similarity).
CC -1- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC family.
DR EMBL; Z32835; CAA83677.1; -.
DR PIR; S44294; S44294.
DR HSP; P04776; 1FXZ.
DR GO; GO:0045735; P:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RnLc like cupin.
DR InterPro; IPR006044; seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PRO0439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S SEED STORAGE; 1.
KW Seed storage protein; Signal; Storage protein.
FT SIGNAL 1 21
FT CHAIN 22 498 legumin A; 11S globulin.
SQ SEQUENCE 498 AA; 56564 MW; 3E75D87D9A9699 CRC64;

Alignment Scores:
Pred. No.: 9,77e-94 Length: 498
Score: 1457.50 Matches: 289
Percent Similarity: 68.88% Conservative: 74
Best Local Similarity: 54.84% Mismatches: 93
Query Match: 54.16% Indels: 71
DB: 2 Gaps: 8

```

```

US-10-728-051-3 (1-1524) x Q41702 (1-498)
QY 1 CGGCAGCAACCGAGAGAGAACCGCTGCCAGTTCCAGCGCTCAATCGCAGAGACCTGAC 60
Db 23 ArgGluGlnSerGlnGlnAsnGluCysGlnLeuGluArgIleAsnAlaLeuGluProAsp 42
QY 61 AATCGCATTAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACCAACGAGGAGTTC 120
Db 43 AsnArgIleGluSerGluGlyLeuIleGluThrTrpAsnProAsnAsnArgGlnPhe 62
QY 121 GAATCGCGCGGCTCCCTCTCTCGCTTAGTCTCCGCGCAACCCCTTCGTAGGCCT 180
Db 63 ArgCysAlaArgValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgArgPro 82
QY 181 TTCTACTCCAATGCTCCCGAGGAGATTCATCCAGCAAGAGGAGGATCTTTGGGTG 240
Db 83 TyrTyrSerAsnAlaProGlnGluIleTyrIleGlnGlnGlyAsnGlyTyrPheGlyMet 102
QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCTTCACACAGAGTCTGTCGATCTCAG 300
Db 103 ValPheProGlyCysProGluThrHisGluGluPro-----GlnGln 116
QY 301 TCCCAAGACCAACCAAGAGAGTCTCCAAAGAGAGAACCAAGCAACAGCAACGAGATAGT 360
Db 117 SerGlu-----GlnGlyGlu-----GlyArgArgTyrArgAspSer 128
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCGACTTCCCAACCGGTGTGCT 420
Db 129 HisGlnLysValAsnArgPheArgGluGlyAspIleIleAlaValProThrGlyIleAla 148
QY 421 TTCTGCTCTACAAACGACACACACTGATGTGTGTGTGTCTTCTTCTTACTGACCAAC 480
Db 149 PheTrpMetTyrAsnAspGlnAspThrProValIleAlaIleSerLeuThrAspThrGly 168
QY 481 AACCAACGACCAACCGCTTCATGATTCCTCCAGGAGATTCATTTGGTGGGAACACGAG 540
Db 169 SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 188
QY 541 CAAGAGTCTTTAAGGTACCAGCAACCAAGCAGACAAAGCAGACGAAAGAGCTTTACCATAT 600
Db 189 GlnGluPheLeuArgTyrGlnHisGln----- 197
QY 601 AGCCCATACAGCCCGCAAGTAGCTAGCCTAGACAAGAGAGCGGTGAATTTAGCCCTCAGGA 660
Db 197 ----- 197
QY 661 CAGCAGCGCGCAGAGACGAGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 198 -----GlnGlyLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 212
QY 721 AGCGGCTTCACGCGCGGAGTTCTCGGAACCAAGCTTCAGGTTTCAGCAGCAGACAGATAGT 780
Db 213 SerGlyPheLysArgAspPheLeuGluAspAlaPheAsnVal---AsnArgHisIleVal 231
QY 781 CAAACCTTAAGAGCGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 232 AspArgLeuGlnGlyArgAsnGluAspGluGluGlyAlaIleValLysValLysGly 251
QY 841 GCGCTCAGATCTTGAGCCCGAGATAGAAAG-----AGACGT 876
Db 252 GlyLeuSerIleIleAlaProGluArgGlnAlaArgHisGluArgGlySerArgGln 271
QY 877 GCGCAGCAAGAGAGAGAGATACGATGAA-----GAT 906
Db 272 GluGluAspGluAspGluLysGluGluArgGlnProSerHisIleLysSerArgArgAsp 291
QY 907 GAATATGAATACGATCAAGAGGATAGA-----AGCGTGGCAGGAGAGAGAGAGAGAGAG 960
Db 292 GluAspGluAspAspLysGluLysArgHisSerGlnLysGlyGlnSerArgArgGlnGly 311
QY 961 GGAATGGTATTGAAGAGAGAGATCTCCAGCGAGTCTGCTAAAAGACATTTGTTAGAAC 1020
Db ----- 1020

```

```

Db      312 AspAsnGlyLeuGluThrValCysThrAlaLysLeuArgAlaAsnIleGlySerSer 331
QY      1021 AGATCCCTGACATCTACAACTTCAAGCTGGTTCACCTCAAACTGCCAACGATCTCAAC 1080
Db      332 ProSerProAspIleTyrAsnProGlnAlaGlyArgIleLysThrValThrSerLeuAsp 351
QY      1081 CTTCTAATACTTAGGTGGCTGGACCTAGTGTGAATATGGAATCTCTACAGGAATGCA 1140
Db      352 LeuProValLeuArgTrpLeuLysLeuSerAlaGluHisGlySerLeuHisLysAsnAla 371
QY      1141 TTGTTTGTCTCCTCACTACACCAACGACGACACAGCATCATATATCGATTGAGGGACGG 1200
Db      372 MetPheValProHisTyrAsnLeuAsnAlaAsnSerValIleTyrAlaLeuLysGlyArg 391
QY      1201 GCTCAGTGCAGTCTGGACAGCAACGACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db      392 AlaArgLeuGlnValValAsnCysAsnGlyAsnThrValPheAspGlyGluLeuGluAla 411
QY      1261 GGTACGTGCTTGTGTGCACAGAACTTCGCCGTGCTGGAAAGTCCACAGACGAGAAC 1320
Db      412 GlyArgAlaLeuThrValProGlnAsnTyrAlaValAlaLysSerLeuSerGluArg 431
QY      1321 TTCGAATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCCAACCTCGCCGTGAA 1380
Db      432 PheThrTyrValAlaPheLysThrAspArgAlaSerIleAlaArgLeuAlaGlyThr 451
QY      1381 AACTCCGTATAGATAACCTGCCGAGGAGGTGTTGCCAAATTCATATGCTCCAAAGG 1440
Db      452 SerSerValIleAspAspLeuProLeuAspValValAlaAlaThrPheAsnMetGlnArg 471
QY      1441 GACGAGGCAAGGAGCTTAAGAACAAACACCCCTTCAAGTTCTTCCTCCACCGTCTCAG 1500
Db      472 AsnGluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProProArgGln 491
QY      1501 CACTCTCCGAGGCTGTGGCT 1521
Db      492 SerGluMetArgAlaSerAla 498

```

Search completed: April 8, 2005, 05:25:50  
Job time : 296.329 secs